

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## The representation of spatial mental models in long-term memory

### Thesis

#### How to cite:

Baguley, Thomas Simon (1994). The representation of spatial mental models in long-term memory. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1993 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000e056>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)



# The representation of spatial mental models in long-term memory

Thomas Simon Baguley  
B.A., Dip. App. Psych. (Nottingham)

July, 1994.

Thesis submitted in part fulfilment  
of the requirements for a Ph.D. in Psychology.

Human Cognition Research Laboratory

Author number: 147085274 and  
date of submission: 28 July 1994  
date of award: 1 November 1994 Psychology,

The Open University,  
Milton Keynes, England.

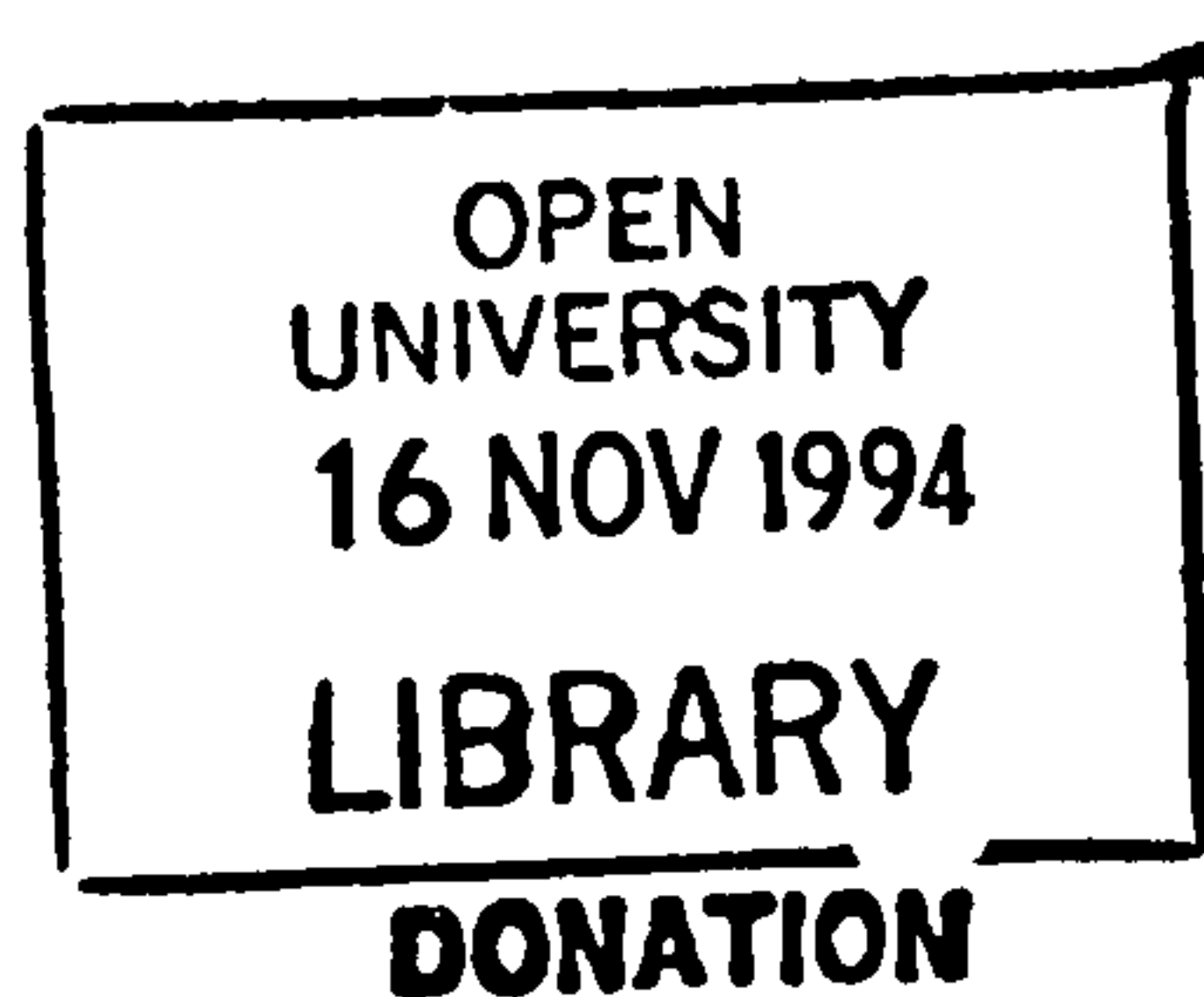
**BEST COPY**

**AVAILABLE**

Variable print quality

# Abstract

This thesis is concerned with how people understand and remember spatial information derived from verbal descriptions. The thesis distinguishes between three different ways of representing spatial information in working memory. The first way is to represent the surface form of the source from which the spatial information is derived (the language of a description). The second is to represent the structure of the situation derived from that source (a spatial mental model). The third is to represent the perceptual characteristics of the situation from a particular perspective (a visual image). Considerable evidence exists that people construct and manipulate spatial mental models in working memory. The purpose of this thesis is to investigate the claim that mental models are represented in long-term memory. An outline of the spatial mental modeling processes required to understand a simple spatial description is proposed. It is proposed that spatial mental modeling is comprised of three processing stages. Firstly, *comprehension* processes are required to access the linguistic meaning of information presented in spatial description. Secondly, *construction* processes are required to build up a representation of the spatial structure of the situation derived from the language of the description. Thirdly, *consultation* processes are required to monitor construction and to access information from the spatial mental model. Nine experiments are reported which investigate evidence for and against the view that people remember the construction and consultation of a spatial mental model. In the final chapter this evidence is reviewed and a 'sketch' of a processing theory of memory for spatial descriptions is proposed. It is argued that memory for a spatial mental model is a product of the interaction between construction and consultation processes over a period of time rather than a simple 'copy' of a completed working memory spatial mental model.



T 153 . 752

Consultation 6py



14 NOV 1994

EX12

HIGHER DEGREES OFFICE

LIBRARY AUTHORISATION FORM

Please return this form to the Higher Degrees Office with the bound library copies of your thesis. All students should complete Part 1. Part 2 applies only to PhD students.

Student: THOMAS SIMON BAGULEY PI: M7085242

Degree: Ph.D.

Thesis title: THE REPRESENTATION OF SPATIAL  
MENTAL MODELS IN LONG-TERM MEMORY

**Part 1 Open University Library Authorisation** (to be completed by all students)

I confirm that I am willing for my thesis to be made available to readers by the Open University Library, and that it may be photocopied, subject to the discretion of the Librarian.

Signed: Thomas Baguley Date: 4/11/94

**Part 2 British Library Authorisation** (to be completed by PhD students only)

If you want a copy of your PhD thesis to be held by the British Library, you must sign a British Doctoral Thesis Agreement Form. You should return it to the Higher Degrees Office with this form and your bound thesis. *You are also required to supply a third, unbound copy of your thesis.* The British Library will use this to make their microfilm copy; it will not be returned. Information on the presentation of the thesis is given in the Agreement Form.

If your thesis is part of a collaborative group project, you will need to obtain the signatures of others involved for the Agreement Form.

The University has agreed that the lodging of your thesis with the British Library should be voluntary. Please tick either (a) or (b) below to indicate your intentions.

(a) ☒ I am willing for the Open University to supply the British Library with a copy of my thesis. A signed Agreement Form and 3 copies of my thesis are attached (two bound as specified in Section 9.4 of the Research Degree Handbook and the third unbound).

(b) ☐ I do not wish the Open University to supply a copy of my thesis to the British Library.

Signed: Thomas Baguley Date: 4/11/94

# Acknowledgements

Special thanks are due to my supervisors Gillian Cohen and George Kiss. To George for getting me started and for clearing up some of my more muddled ideas. To Jill for her patience and hard work and most importantly her ability to tread the careful path between criticism and encouragement. Thanks also to Steve Payne for several productive conversations on memory and mental models.

I would like to thank all those in Psychology and H.C.R.L. who have made writing this thesis more of a pleasure than a chore (I think). This, in turn, is in no small part a vote of thanks to Marc Eisenstadt and Richard Stevens in their respective official roles as Director of H.C.R.L. and acting head of the Psychology discipline.

I would also like to include my family and friends and, of course the extended family of The Open University. In particular I ought to mention Birgit, Dot, Emma, Hank, Ian, Ingrid, Jonathan, Kathryn and Simon.

Finally, I would like to thank all those Open University members of staff and students who gave up some of their time to take part in one or more of my experiments.

This research was supported by a studentship from the Science and Engineering Research Council and by the Human Cognition Research Laboratory and Faculty of Social Science of the Open University.

# Open University Higher Degrees Requirements

## Statement in association with rule HD28:

The work described in this thesis is entirely the work of the candidate unless otherwise stated and has not been submitted elsewhere for any other qualification.

## Statement in association with rule HD28:

Material from this thesis has been presented in two conference papers. The first paper was based on Experiments 1 and 2 reported in Chapter 3:

Baguley T.S. (1993). Ordinal recall in spatial mental models. Paper presented at the *Sixth Conference of the European Cognitive Psychology Society*. September 11-15, Elsinore, Copenhagen, Denmark.

The second paper was based on Experiments 4 and 6b reported in Chapter 4:

Baguley T.S. (1994). Are spatial mental models represented in long-term memory? Paper presented at the *Annual Conference of the British Psychological Society*. March 24-27, Brighton, England.



# Table of Contents

<b>Chapter 1 - Representation, memory and space.....</b>	<b>1</b>
1.1 A functional approach to representing, understanding and remembering spatial situations.....	2
1.2 The role of representation in cognition.....	3
1.2.1 Perception.....	4
1.2.2 Action .....	5
1.2.3 Language.....	6
1.2.4 Reasoning.....	7
1.3 Representation and memory.....	8
1.3.1 Representation and process.....	8
1.3.2 The imagery debate .....	10
1.3.3 Levels of representation.....	14
1.4 The perceptual and structural characteristics of imagery.....	16
1.5 Memory for objects, scenes and events.....	18
1.6 The influence of working memory processes on long-term memory .....	21
1.6.1 Memory and encoding processes .....	21
1.6.2 The influence of retrieval cues on recall.....	22
1.7 Conclusions.....	24

<b>Chapter 2 - The representation of spatial mental models in memory.....</b>	<b>25</b>
2.1 The theory of mental models.....	26
2.1.1 What is a mental model? .....	26
2.1.2 Mental models and reasoning.....	27
2.1.3 Mental models and discourse comprehension .....	28
2.1.4 The characteristics of mental models .....	29
2.1.5 Alternative conceptions of mental models.....	33
2.1.6 Mental models as representations of space.....	33
2.1.7 Mental models and spatial reasoning .....	35
2.2 Mental models and theories of working memory.....	37
2.2.1 A specific model of working memory.....	37
2.2.2 Mental models and specific working memory systems.....	39
2.3 Experimental studies of mental models in working memory.....	41
2.4 Mental models and long-term memory.....	43
2.4.1 Memory and reasoning with mental models .....	43
2.4.2 Mental models and memory for text .....	44

2.4.3 Mental models and the fan effect .....	45
2.4.4 Referential continuity and the construction of spatial mental models .....	46
2.4.5 The determinacy of spatial descriptions.....	48
2.4.6 Perspective, order and memory for spatial mental models .....	51
2.5 What constitutes memory for a spatial mental model?.....	53
2.5.1 An outline of spatial mental modeling.....	53
2.5.2 How many ways can a spatial mental model be remembered?.....	57
2.5.3 Conclusions.....	60

## **Chapter 3 - Ordinal recall as evidence of memory for a spatial mental model.....61**

3.1 Challenges to Tversky's account of ordinal effects in drawing recall .....	61
3.2 Preface to the experimentation.....	63
3.3 Experiment 1 .....	64
3.3.1 Method.....	65
3.3.2 Results.....	68
3.3.3 Discussion.....	75
3.4 Experiment 2 .....	78
3.4.1 Method.....	81
3.4.2 Results.....	83
3.4.3 Discussion.....	88
3.5 Experiment 3 .....	90
3.5.1 Method.....	91
3.5.2 Results.....	94
3.5.3 Discussion.....	100
3.6 General Discussion .....	103
3.6.1 Drawing orders reflect a spatial mental model rather than an image.....	105
3.6.2 Comprehending scrambled texts requires a mental model.....	106
3.7 Conclusions .....	106

## **Chapter 4 - Determinacy and memory for spatial descriptions.....108**

4.1 Experiment 4 .....	111
4.1.1 Method.....	114
4.1.2 Results.....	116
4.1.3 Discussion.....	126
4.2 Experiment 5 .....	129
4.2.1 Method.....	130



4.2.2 Results .....	130
4.2.3 Discussion .....	139
4.3 Experiment 6a.....	142
4.3.1 Method .....	142
4.3.2 Results .....	143
4.3.3 Discussion .....	146
4.4 Experiment 6b .....	147
4.4.1 Method .....	147
4.4.2 Results .....	149
4.4.3 Discussion .....	156
4.5 General Discussion.....	159
4.5.1 Memory for the language of a description .....	159
4.5.2 Memory for an episodic construction trace .....	160
4.5.3 Memory for the structure of a spatial mental model.....	162
4.6 Conclusions.....	162

## **Chapter 5 - The long-term retention of spatial information.... 164**

5.1 Very long term memory for pictures and places .....	165
5.2 Recall, reconstruction or guesswork.....	166
5.3 Very long term memory for spatial mental models .....	167
5.4 Experiment 7 .....	168
5.4.1 Method .....	169
5.4.2 Results .....	171
5.4.3 Discussion .....	177
5.5 Experiment 8.....	179
5.5.1 A brief description of the original task .....	179
5.5.2 The experiment.....	180
5.5.3 Method .....	181
5.5.4 Results .....	182
5.5.5 Discussion .....	191
5.6 Experiment 9.....	194
5.6.1 Method .....	195
5.6.2 Results .....	196
5.6.3 Discussion .....	207
5.7 General Discussion.....	209
5.7.1 Assessing level and type of recall.....	209
5.7.2 What kinds of information are preserved? .....	210
5.8 Conclusions.....	211

**Chapter 6 - Conclusions.....212**

6.1 Spatial mental models and memory for spatial descriptions .....213

6.1.1 Are drawing orders influenced by the order of spatial mental model construction? .....214

6.1.2 Is the probability of recall in a drawing task influenced by the order of spatial mental model construction? .....215

6.1.3 Does reordering a spatial description impair recognition memory?.....216

6.1.4 Do people remember the structure of a spatial mental model?.....217

6.1.5 Discussion.....218

6.2 A processing account of the representation of spatial mental models in long-term memory .....220

6.2.1 Copy theories of memory for spatial mental models .....220

6.2.2 A sketch of a processing theory of memory for spatial mental models .....221

6.3 Theoretical perspectives on memory for spatial mental models .....231

6.4 Methodological issues.....233

6.4.1 Distinguishing between different forms of mental representation .....234

6.4.2 Distinguishing between working memory and long-term memory representations.....234

6.4.3 Ecological validity.....235

6.4.4 Strategies and individual differences.....236

6.5 Representational issues.....237

6.5.1 The relationship between images and mental models .....237

6.5.2 Schemas and mental models in long-term memory.....238

6.5.3 The format of spatial mental models in long-term memory.....239

6.6 Implications for future research .....240

**References.....242**



# Chapter 1 - Representation, memory and space

Mental representations denote or stand for entities, concepts and relationships in the world around us. In order to understand the world we need to construct a *representing world* which is a simplification, and sometimes a distortion, of the 'real' *represented world* (Johnson-Laird, 1983; Newell, 1990; Palmer, 1978; Rumelhart & Norman, 1983). Stephen Palmer (1978) has stressed the importance of the relationship between how things are represented and what procedures are able to act upon those representations. Drawing upon this idea it is possible to distinguish between a number of different high-level representations; representations of verbally expressible propositions, representations of structural relations of real or imagined situations and representations of the perceptual characteristics of situations. Using terminology proposed by Philip Johnson-Laird (1983) these would correspond to *linguistic propositions*, *mental models* and *images*. A particularly important distinction proposed here is between images which preserve structural relations among parts of the image (and which thus imply the existence of a mental model) and those which do not (and hence do not imply the existence of a mental model - merely the experience of perceptual or affective imagery). This issue is explored further in section 1.4. Spatial relations are one of the most fundamental, if not the most fundamental, example of the structural relations that the human mind is capable of representing and understanding. Mental models of the structure of spatial situations are referred to as *spatial mental models*. People need to be able to construct spatial mental models in order to understand their physical environment and to interact with it. This can be demonstrated by looking at the way spatial representations and particularly spatial mental models are used in human cognition (such a functional approach is described in section 1.2).

There is considerable evidence that spatial mental models are constructed and manipulated in working memory (this evidence is reviewed in Chapter 2). The purpose of this thesis is to investigate whether spatial mental models are represented in long-term memory.

Later in this chapter the aspects of representation and memory raised in the previous paragraph are discussed in more detail. The purpose of this discussion is to show how high-level working memory representations can support qualitatively different kinds of processing. These qualitative differences make it possible to distinguish between different forms of representation and also influence what information is retained in long-term memory. The durability and accessibility of memory traces in long-term memory is greatly influenced by encoding processes in working memory; for example by the overlap between information encoded onto the memory trace and information present in the retrieval context (see sections 1.6.1 and 1.6.2). For this reason (and because of the intimate relationship between representation and process discussed later in this chapter) investigating the representation of spatial mental models in long-term memory requires an understanding of the representation of mental models in working memory and of the processes which act upon them.

Chapter 1 introduces a functional approach to representation and memory, with particular attention on the role of spatial representation in cognition. Chapter 2 will discuss the representation of mental models in working memory and review evidence of their retention in long-term memory. Chapter 2 concludes with a proposal of an outline of how spatial mental models are constructed and manipulated in working memory. Chapters 3, 4 and 5 present empirical evidence about long-term memory for spatial information derived from verbal descriptions. Chapter 6 discusses the implications of the research presented in earlier chapters and presents a 'sketch' of a theory of long-term memory for spatial mental models. Chapter 6 ends with a discussion of some of the important issues raised by the thesis and presents some conclusions based on the empirical and theoretical work presented here.

## **1.1 A functional approach to representing, understanding and remembering spatial situations**

What kind of properties does our understanding of and memory for spatial situations require of the way we represent spatial information? The most important property is the capacity to preserve the spatial structure present in the world. Consider three different ways in which people construct external representations of spatial situations. One way would be to write a description of the spatial situation made up of sentences such as "The chair is to the left of the desk". A second way would be to draw a picture, or better still take a photograph, of the spatial situation. A third way would be to draw a diagram or make a scale model of the situation. All three of these ways are common methods of recording spatial information in everyday life, and each has advantages and disadvantages. Verbal descriptions can be very economical. There is no need to describe every possible spatial relation present in the situation, certain key relations between objects or places can often suffice to reconstruct or capture the relevant details of the situation. By the same token verbal descriptions also carry a degree of uncertainty. For example.



if someone says that “The chair is to the left of the desk” it is unclear whether the chair is one inch to the left or one yard to the left. A picture or photograph does not have this problem of uncertainty. Pictures and photographs, however, capture the spatial characteristics only indirectly. What they really preserve are perceptual characteristics of the situation from a particular perspective. Deriving spatial relations from a picture or photograph is not always easy, because of this problem of perspective. The problem of perspective can be avoided if a diagram or scale model is employed. A diagram or scale model in some sense preserves the spatial structure of the situation it represents (Craik, 1943). At the same time, though, other information may be lost. Typically, diagrams and scale models don’t preserve much in the way of perceptual detail. Often they are at a level of abstraction which ignores surface features such as colour or texture. Diagrams, and particularly scale models, also tend to be difficult to construct.

There are strong similarities between these three ways of constructing external representations of space and the tripartite distinction between linguistic propositions, images and mental models adopted elsewhere in this thesis (Johnson-Laird, 1983). It is not proposed that the use of these kinds of external representation provides direct evidence of this distinction in the way people represent spatial situations in working memory or long-term memory. The way people use and interact with external representation to some extent ‘maps out’ the possible ways of representing their physical environment. Any theory of mental representation must be able to account for the whole range of ways people interact with their environment. What this probably means is that people make use of a number of different forms of spatial representation for a variety of different purposes. Further evidence for this is provided by looking at the different purposes for which spatial information is required. The role of spatial representation in cognition is illustrated with respect to four major areas of psychological research in the following section.

## **1.2 The role of representation in cognition**

Work on representation has fallen primarily into one of four broad areas: perception, action, language and reasoning (Kosslyn, 1984). In the following four sections the function of representation in each of these areas is illustrated with examples taken from the literature in experimental psychology and cognitive science. The emphasis, in each example, is on spatial representation, because this is the focus of this thesis and because studies of spatial representation have provided the most striking evidence that people take advantage of multiple forms of representation.



### 1.2.1 Perception

A fundamental problem in perception is the question of how people are able to recognize a three-dimensional object from a two-dimensional visual image (Marr, 1980; 1982; Roth & Frisby, 1986). One way of accomplishing this is for people to store information about the three-dimensional structure of an object in long-term memory and to compare new patterns against this stored representation. *Structural descriptions* are modality free, abstract representations of the information necessary to construct the original object (Palmer, 1975; 1978). Reed (1974) has provided evidence that pattern recognition involves structural descriptions rather than visual images or fixed templates. Reed presented subjects with two geometrical patterns in sequence. Subjects were asked to decide whether the second pattern formed part of the first pattern. Repeating a whole pattern resulted in the greatest number of correct responses. There were also very large differences in accuracy for different parts of the original pattern. Some parts were almost six times harder to recognize than others. This is not consistent with the use of a fixed template for recognition - a template should not show marked differences in recognition for different parts of the patterns (Reed, 1974; Roth & Frisby, 1986). Reed obtained similar results for long 5 second retention intervals as for short 1.5 second intervals which he argued would preclude subjects from scanning a visual image. They are consistent with a representation made up of parts and relations between parts.

The importance of Reed's findings are that they stress the importance of structural information in perception and representation. A similar, but weaker, point of view can be found in work by Shepard and Cooper on mental rotation (Shepard & Cooper, 1982; Shepard & Metzler, 1971). The 'images' described by Shepard "are highly abstract and schematic in comparison to a concrete perceptual image but preserve enough of the essential structure to permit accurate comparisons in their experiments" (Reed, 1974). Structural descriptions are usually considered to be propositional representations, however this is not necessarily the case (Reed, 1974; Roth & Frisby, 1986). Any representation which preserves the structure of the original object and which supports perceptual processes which can be used to generate two-dimensional views or recognize an object can be considered as a structural description. For example an architects' scale model of a house could be considered a structural description just as much as a set of equations encoded in a computer design program (Palmer, 1978; Tversky, 1991). There is a crucial distinction between representation of abstract structure and representations of concrete, perceptual characteristics. Perception requires both, but without a structural representation it is difficult to account for the recognition of familiar objects from unfamiliar views or perspectives (Johnson-Laird, 1989; Marr, 1982; Roth & Frisby, 1986).



### 1.2.2 Action

Representational accounts of action are an important area in which cognitive psychology has challenged behaviourism. For example, Tolman (1932; 1948) argued that stimulus-response accounts were inadequate to explain rats ability to generalize action from dry to water-filled mazes. He argued that navigation requires the retention of knowledge in the form of a cognitive map which is independent of the precise motor behaviour of an organism. Despite importance of representational issues to an understanding of action it is a much neglected area of cognitive psychology (Kelso, 1982; Rosenbaum, 1984). One partial exception has been the study of navigation in animals and in humans. Navigation requires the coordination of spatial representations with the external, physical environment. Of particular interest are the differences in spatial representations acquired from actual navigation in an environment and those acquired from other sources. Thorndyke and Hayes-Roth (1982) compared spatial knowledge acquired from learning a map (*survey* knowledge) or from navigation (*route* knowledge). They tested spatial memory of the Rand Corporation building. Subjects who had learned from a floor plan of the building were better at judgements of relative location and straight-line distance. Subjects who learned the layout of the building from actual navigation were better at estimating route distances and orienting themselves relative to unseen locations. With further experience of navigation the advantages for map learning subjects on location and distance memory tests disappeared.

Thorndyke and Hayes-Roth took this as evidence for the representation of two types of spatial information: *survey* and *procedural* knowledge. Map learning provides survey knowledge of global spatial relations in the form of visual or spatial images. Navigation supports the learning of procedural knowledge of specific routes. This procedural knowledge is sequential and may be based on location specific views of the environment. It can be consulted by 'mentally simulating' navigation along the route. They also found that subjects who had considerable experience navigating the building demonstrated accurate survey knowledge almost as if they could see through the walls of the building. What is particularly fascinating about this study is the way people are able to integrate procedural, route knowledge over a period of time in such a way that it can mimic or even exceed accurate survey knowledge. Thorndyke and Hayes-Roth describe it as possessing the characteristics of a three-dimensional model of the environment. This representation can only be the result of combining or integrating route knowledge learned on different occasions. This in turn implies that individual routes are, to some extent, able to preserve the spatial structure of the environment. Experiments with small-scale spatial situations have shown similar results, with route learning culminating in an integrated representation of an environment (Levine, Jankovic, & Palij, 1982). It appears that flexible, detailed structural representations of the environment are a natural consequence of action in the real world.



### 1.2.3 Language

Chomsky challenged behaviourist explanations of language with representational accounts of grammar, citing language acquisition and the ability to generate novel grammatical utterances as evidence (Gardner, 1987; Leahey, 1992). Cognitive psychologists have looked at issues of broader interest in cognition. These include sentence memory and language comprehension. For example, it has been proposed that people use the explicit content of discourse only as a basis for comprehension, filling in for and making inferences about missing details as necessary (Cohen, 1989; Garnham, 1987; Johnson-Laird, 1981; van Dijk & Kintsch, 1983). In a famous study Bransford, Barclay and Franks (1972) gave subjects a number of sentences constructed from the same basic sentence frame:

“Three turtles rested (on/beside) a floating log and a fish swam beneath (it/them)”

Bransford and his colleagues showed that subjects presented with sentences such as “*Three turtles rested on a floating log and a fish swam beneath them*” would later confuse them with sentences of the form “*Three turtles rested on a floating log and a fish swam beneath it*”. Bransford, Barclay and Franks argued that subjects were not simply representing the semantic deep structure of the sentences they had learned, but were abstracting (or constructing) representations of the situations being described. They noted that, because of the nature of the stimuli used in their experiments, people may generate spatial representations of the situations described, and these may take the form of mental images. Constructive processes in sentence comprehension result in semantic representations, though not necessarily in the form of mental images. They proposed that representations of the language alone are not sufficient to explain language comprehension and that semantic representations of situations play a functional role in language understanding (Bransford et al., 1972; Bransford & Franks, 1971; Garnham, 1981; Garnham, Oakhill, & Johnson-Laird, 1982). More recent psychologists and psycholinguists would refer to these semantic representations as mental models, discourse models or situation models. Despite the confusing mix of terminology they would agree on the main conclusion of Bransford and his colleagues that comprehension involves the construction of a coherent representation of the meaning or situation described by discourse (Albrecht & O’Brien, 1993; Bower & Morrow, 1990; Garnham, 1987; Glenberg, Meyer, & Lindberg, 1987; Johnson-Laird, 1983; Schmalhofer & Glavanov, 1986; van Dijk & Kintsch, 1983). Understanding how people use language to convey spatial information is therefore an important element of any theory of spatial representation. The analysis of how spatial information is expressed through language has a long tradition in psychology and linguistics (Clark, 1973; Levelt, 1984; Miller & Johnson-Laird, 1976). However, in this thesis, spatial language will be considered only as a starting point for or an input to a spatial representation, rather than as an area of study in its own right.



## 1.2.4 Reasoning

Traditional theories of reasoning derived from philosophy and mathematics propose that people make deductive inferences by the application of formal, logical rules (Braine, 1978; Henle, 1962; Piaget & Inhelder, 1958; Rips, 1983). One of the strongest challenges to this view has been found in the area of relational reasoning. In fact the account of relational reasoning described here can be traced back at least as far as William James or even Aristotle (Garnham & Oakhill, 1994; Hunter, 1957; Johnson-Laird & Byrne, 1991). Relational inferences are made from premises which can be expressed on a linear dimension such as “A is better than B” or “X is above Y”. These premises lend themselves to being represented spatially. One plausible alternative to logic-based accounts is that people construct an integrated representation of the premises and use this rather than a set of formal rules to derive a conclusion. Most versions of this account have assumed that this integrated representation is spatial or visual, but the essential feature of the explanation is that all the information in the premises is integrated into a single coherent representation during reasoning. Ian Hunter (1957) proposed that people transform the premises of relational problems in order to construct an integrated representation in primary or working memory. Imagine being presented with a three-term series problem in the form:

Ingrid is sadder than Dorothy.

Emma is happier than Dorothy.

Who is the happiest?

According to Hunter people could solve this problem using two kinds of operation. *Conversion* of the first premise (reversing the relation) would produce “Dorothy is happier than Ingrid”. *Reordering* of the two premises results in a representation where the two middle terms are adjacent. Evidence that people use these two operations to construct an integrated representation of the premises according to the exact rules proposed by Hunter is mixed. However, evidence that these or similar operations are used in relational reasoning is very strong (Evans, Newstead, & Byrne, 1993). Problems which require premise integration are harder than those, like two-term series problems, which do not (Clark, 1969), while five-term series and three-term series, both of which involve premise integration, are equally difficult (Byrne & Johnson-Laird, 1989). Problems with premises which are continuous, and hence which do not need to be transformed, are easier than those which are discontinuous (Ehrlich & Johnson-Laird, 1982; Potts, 1972). These issues are discussed in more detail in a review by Evans, Newstead and Byrne (1993) who conclude that relational reasoning does involve an integrated representation of the premises. The idea that operations are carried out in working memory to transform verbal descriptions into integrated representations lies at the heart of reasoning theories based on mental models. Premise conversion, in particular, is implicit in the mental model construction processes described by Johnson-Laird (1983) and Payne (1993). An explicit account of why and how verbal descriptions are converted during spatial mental model construction is provided at the end of Chapter 2.



## 1.3 Representation and memory

In the rest of this chapter the concept of mental representation will be discussed in relation to memory research in cognitive psychology. The concept of representation is central to information processing theories (Fodor & Pylyshyn, 1988; Palmer & Kimchi, 1986; Pylyshyn, 1984; Rumelhart & Norman, 1983). This is because the origins of cognitive psychology lie in the necessity to postulate knowledge that is held by an organism which is independent of the way that knowledge influences behaviour (Tolman, 1932; 1948). This knowledge is held to be essential to explain complex behaviour in areas such as language comprehension, problem solving, perception and action (as illustrated earlier in section 1.2). Many arguments and debates in psychology arise because of different interpretations or definitions of common constructs such as representation (Broadbent, 1985; Kosslyn & Pomerantz, 1977; Palmer, 1978; Pylyshyn, 1973; Rumelhart, 1989; Smolensky, 1988). It is important to establish the assumptions cognitive psychologists share when they write or talk about representation. For the same reasons genuine differences also need to be outlined. Theoretical accounts of memory, and in particular human memory, are examined for two reasons. Firstly, because the development of accounts of different kinds of mental representation are intimately linked to the history of memory research in psychology and neuroscience (Anderson & Bower, 1973; Hintzman, 1990; McClelland & Rumelhart, 1981; Paivio, 1971). Secondly, this thesis examines evidence that a particular form of mental representation in working memory is also represented in long-term memory (Glenberg et al., 1987; Mani & Johnson-Laird, 1982; Payne, 1993). This chapter provides a general introduction to issues surrounding the relationship between representation, working memory and long-term memory. Subsequent chapters discuss the nature of mental models and present original empirical evidence that supports the thesis that people construct and retain spatial mental models in long-term memory.

### 1.3.1 Representation and process

The simplest definition of a representation is that it is an entity that stands for, or denotes, another entity. This definition is something that most, if not all, cognitive psychologists would agree with (Eysenck & Keane, 1990; Palmer, 1978; Rumelhart & Norman, 1983). A representation can be an external entity such as a portrait, flag or emblem, or it can be mental entity, for example an idea, an association or an image. Many cognitive psychologists believe that mental representations are somehow related to the 'phenomenological' experience of mental events and play a causal or mediating role in behaviour (Bruce & Green, 1985; Marr, 1982; Pylyshyn, 1984; Rumelhart & McClelland, 1985). Cognitive psychologists are primarily concerned with mental representation. However, many psychologists are interested in external representations, either in themselves, or for what they can tell us about the mind. Of particular interest are the



representations that are difficult to classify as internal or external (e.g. the written, spoken or signed word).

Some cognitive scientists, psychologists and philosophers place a stronger interpretation on what it means for something to be representational. They argue that mental representations are arbitrary symbols with particular formal properties (Chomsky, 1980; Fodor, 1975; Fodor, 1980). According to Fodor (1980) formal properties "... are specified without reference to such semantic properties of representations as, for example, truth, reference, and meaning." An example would be a structural property such as syntax. This strong 'symbolic' or 'computational' view, can be regarded, at least until the advent of connectionism, as the dominant approach in cognitive science, particularly in artificial intelligence, linguistics and cognitive psychology (Hunt, 1989; Rips, 1986). Dominant is not taken here to imply the consensus or even majority view. Even before the growth in popularity of the 'new' connectionism many cognitive scientists placed a different interpretation on computation and representation (Johnson-Laird, 1983; Kosslyn & Pomerantz, 1977; Norman & Rumelhart, 1975; Paivio, 1971). In this thesis the weaker, more general view of what constitutes representation in information processing psychology is adopted (Johnson-Laird, 1983; Palmer, 1978; Palmer & Kimchi, 1986; Rumelhart & Norman, 1983). According to this view representations are considered to have constituent structure, but this structure is not necessarily based purely on formal 'syntactic' properties. This is the view that "different types of representation are logically distinguishable at some level of analysis, and, moreover, that they exist as different options for encoding information" (Johnson-Laird, 1983). It is a level of analysis at which cognitive processes act upon representations and at which it is meaningful to construct theories which can be tested by traditional experimental psychological procedures.

A major problem for theories of representation has been how to elucidate the correspondence between the representing world and the represented world. This problem lies at the heart of a number of controversies in psychology and cognitive science (Broadbent, 1985; Kosslyn & Pomerantz, 1977; Pylyshyn, 1973; Rumelhart & McClelland, 1985; Smolensky, 1988). Palmer (1978) illustrates this by arguing that a representation should not be considered as a single entity but as part of a representational system; linking two "functionally separate worlds" which are related in some way. A representation can be specified by stating the relevant aspects of the *represented world* that refer to the *representing world*. Palmer (1978) argues that to do this requires a statement of:

- (1) what the represented world consists of,
- (2) what the representing world consists of,
- (3) the aspects of the represented world being modeled,
- (4) the aspects of the representing world doing the modeling, and
- (5) the correspondences between the two worlds.

Rumelhart and Norman (1983) express the same idea in terms of sets. The represented world and the representing world are made up of a set of objects and relations among those objects. The representing world is mapped onto a subset of the represented world (the *to-be-represented world*). The two important correspondences are thus between the objects and the relations of this *to-be-represented world* and the objects and relations of the *representing world*.

These descriptions both place emphasis on the representational systems, rather than on individual representations per se. Palmer (1978) argues that these processes define or limit the informational content of the representation. Process issues are thus also central to mental representation. In practice representation and process are very difficult to separate because mental events are not directly observable. The notion that process and representation are fundamentally linked in any representational system is a central feature of the model of memory for spatial mental model construction and consultation outlined in this thesis (Chapters 2 and 6).

### 1.3.2 The imagery debate

Psychology and cognitive science have been dominated by a number of controversies and debates relating to the nature of mental representation (Gardner, 1987; Leahey, 1991). Early psychologists debated questions such as whether it was possible to have 'imageless' thought (Boring, 1950; Dellarosa, 1988; Leahey, 1991). Behaviourists rejected the notion that knowledge could be represented independently of the way it was used (Boring, 1950; Tolman, 1932; Tolman, 1948; Watson, 1913). In recent years the ground has shifted to a slightly different kind of question: is it necessary to postulate more than one form of representation in working memory or in long-term memory? Possibly the most important debate in relation to this question has concerned the nature of mental imagery. The debate can be divided roughly into two phases. In the early phase the controversy concerned whether it was possible or reasonable to distinguish between imagery and verbal representations as distinct long-term memory codes. This phase was dominated by the work of Allan Paivio on memory and imagery using a traditional verbal learning methodology (Paivio, 1971; 1986; 1991). The second phase was dominated by striking empirical observations that the process of inspecting a visual image was very similar to perceiving a real picture or object (Kosslyn, Ball, & Reiser, 1978; Shepard & Metzler, 1971).



## The rise and fall of dual-coding theory

When memory for different stimuli is compared pictures are consistently remembered better than words (Denis, 1991; Eysenck, 1984; Glenberg & Langston, 1992; Intraub, 1979). The most influential explanation of this picture superiority effect is *dual-coding theory* (Paivio, 1971; 1986; 1991). Dual-coding theory proposes two independent, but interacting, symbolic systems; an image system and a verbal system. The image system is specialized for dealing with perceptual information, while the verbal system is specialized for dealing with linguistic information. While distinct encoding and retrieval systems are associated with each system the picture superiority effect arises because two distinct codes are stored for pictures (an imaginal code and a verbal label) whereas under normal circumstances words are only stored using a verbal code. According to dual-coding theory when imagery instructions are given or when people spontaneously adopt an imagery strategy (e.g. when people learn lists of concrete, easily imaged words) the picture superiority effect is eliminated or reduced (Paivio, 1971; 1986; 1991).

Dual-coding theory has come under attack from two main directions. The first line of attack is in relation to the nature of mental representation. Propositional theorists in the imagery debate have attacked the idea that imagery plays any causal role in cognition (Pylyshyn, 1973; 1984). The second line of attack comes from theorists who are willing to accept the view that imagery can be meaningfully studied and that it may play a causal role in cognition. It is this second line of attack that has been most damaging to dual-coding theory. These imagery theorists propose that there is a separate, functional imagery system in working memory, but that there is no functionally distinct long-term memory imagery code (Anderson & Bower, 1973; Denis, 1991; Hampson & Morris, 1979; Kosslyn, 1980; Marschark & Surian, 1989; Nelson, 1979; Yuille & Catchpole, 1977). Evidence has accumulated against dual-coding explanations of picture superiority and concreteness effects from a number of different sources.

Research on the influence of different imagery instructions has shown that it is relational or linking imagery rather than imagery per se that improves recall for concrete words. Item specific imagery (concentrating on the visual features of the imagined entity) or separation imagery (imagining two or more spatially separated entities) do not improve memory for concrete words whereas relational imagery (imagining two or more entities in some form of visual or spatial relationship) does. These studies have supported a number of processing or encoding accounts of the effect of imagery on memory (Bower, 1970; Hunt & Einstein, 1980; Marschark, 1985; Marschark & Surian, 1989; Morris & Stevens, 1974; Morris and Reid, 1975). A related finding is that the influence of imagery on memorability appears to be restricted to the visual mode. Ellis (1991) has demonstrated that neither olfaction, audition or touch show the same 'imagery' effects as vision in intentional, incidental or paired-associate learning as well as on an imagery Stroop task. This goes against coding redundancy accounts favoured by Paivio (1991), but



is consistent with the relational processing account of Marschark and Surian (1989). Ellis himself explains his findings in terms of the *parallelism* of visual perception and visual imagery which supports relational processing in a way that other modalities do not (Ellis, 1991). These findings leave open the possibility that other modalities may support different kinds of processing during encoding.

Other evidence is also problematic for dual-coding theory. The theory assumes that picture are always ‘implicitly’ named. Intraub (1979) has shown that naming latency does not influence the retention of pictures. Nelson (1979) has shown when pictures are not implicitly named the picture superiority effect is not reduced or eliminated. Neurophysiological and neuropsychological studies have suggested that there is a functional distinction between spatial and visual aspects of imagery, complicating the conception of a single visual imagery system (Farah, 1988; Farah, Hammond, Levine, & Calvanio, 1988). At least one prominent neuropsychologists has concluded that such case studies do not reflect functionally distinct systems of *storage* in semantic memory, in fact only a handful of these deficits are thought to implicate semantic memory rather than input or output processes (Shallice, 1988). Weldon and Roediger (1987) have noted that the retention of words surpasses that of pictures in data-driven implicit memory tests and have proposed an explanation of picture superiority and other imagery results based on the notion of transfer-appropriate processing (Bransford, Franks, Morris, & Stein, 1979; Morris, Bransford, & Franks, 1977).

While a great deal of experimental evidence has accumulated against dual-coding theory it has still retained some popularity. Paivio has also defended dual-coding theory against many of the findings presented here, however many of his responses have emphasised encoding and retrieval processes such as the “conceptual peg hypothesis” at the expense of coding redundancy (Paivio, 1991). In conclusion, there is relatively little evidence to support the strong assumption that functionally distinct imaginal and verbal codes are represented and stored in long-term memory. Alternative explanations have focused on characteristics of the retrieval situation or, perhaps more often, on encoding processes (see section 1.6).

### **How are images represented in working memory?**

Imagery theorists such as Kosslyn and Shepard rejected the notion that there are separate imaginal and verbal codes in long-term memory (Kosslyn, 1980; Kosslyn, 1981; Shepard & Cooper, 1982; Shepard & Podgorny, 1978). Kosslyn in particular stressed the distinction between structural codes stored propositionally in long-term memory and special-purpose working memory representations temporarily constructed and manipulated in a spatial medium. His mental scanning experiments have provided evidence that inspecting an image is similar to inspecting a visual scene, with size, granularity and distance of the imaged object all having appropriate effects (Kosslyn, 1981; Kosslyn et al., 1978). For example it is easier to ‘see’



information in a large image than a small image, and the time taken to scan across a visual image increases with the distance between the scanned parts of the image. Shepard and colleagues have shown that mental rotation of two and three dimensional objects is compellingly like actually perceiving such a rotation (Shepard & Cooper, 1982; Shepard & Feng, 1972; Shepard & Metzler, 1971). Shepard, Kosslyn and others argued that imagery was not dependent on the manipulation of arbitrary propositions but on *analogical* representations which are “quasi-pictorial” in nature. The mental processes involved are similar to those underlying the perception of a picture or a visual scene (Johnson-Laird, 1983; McGuinness, 1989). Critics of imagery like Zenon Pylyshyn have argued that imagery is epiphenomenal. In other words that the image or the experience of the image plays no causal role in cognitive activity. Despite the wealth of experimental evidence on imagery the distinction between analogical and propositional representations is still controversial. The central question, therefore, is what does it entail for a representation to be analogical?

### Digital or analogical representation?

One possible candidate is that analogical corresponds to continuous representation and propositional to digital representation. Shepard and Cooper’s experiments have stressed the importance of demonstrating that mental rotations are continuous in that they progress through intermediate transformational states. Cooper (1976) measured rates of mental rotation for subjects. She was able to demonstrate a linear relationship between angular departure from the expected orientation of the object and time to decide whether it was a mirror image of the test item or not. However, it is a misunderstanding to suggest that the importance of analogical representations is that they are continuous. Mental rotation *must* be continuous in order to mimic or approximate to an actual rotation (Palmer, 1978; Rumelhart & Norman, 1983). A continuous ‘analogical’ representation of a digital process would be a poor approximation of the original process, though it may be that analogical representations may, for other reasons, be more appropriate for modeling continuously varying information (Rumelhart & Norman, 1983).

### Intrinsic or extrinsic representation?

Palmer (1978) proposes a distinction between intrinsic and extrinsic representation. Analogical theorists (Kosslyn, 1980; Paivio, 1986) argue that not only do images represent spatial dimensions, but they are represented in a spatial medium such that the image resembles what it represents. According to Palmer (1978; Shepard, 1975; Shepard & Chipman, 1970) this claim is weaker than *physical isomorphism* (e.g. the ‘picture metaphor’ of imagery), but stronger than *functional isomorphism* where the representation maps changes in the real world directly onto objects in the represented world (e.g. a physical and mental rotation that both pass through intervening stages). Intrinsic (analogical) representations use relations that have the same inherent structure as the relations being modeled. Extrinsic representations possess only arbitrary, syntactic structure, in that they do not attempt to approximate to the structure of the



represented situation. Therefore, for an extrinsic representation the relevant relations that exist in the real world have to be explicitly built into the system by reference to the external world (Palmer, 1978; Rumelhart & Norman, 1983).

### Attempts to resolve the imagery debate

Pylyshyn (1973; 1984) argues that imagery requires tacit knowledge of what it would be like to perceive the stimuli, rather than relying on similarities between the processes of perception and imagery. This tacit knowledge (which is not accessible to conscious experience) of visual processes and the visual world is used to generate information about what would be seen if a particular object or scene were present. The appeal to tacit knowledge is problematic because it is difficult and probably impossible to falsify. Pylyshyn has asserted that images are 'cognitively penetrable' because they can be influenced by propositional desires, beliefs or goals (Boden, 1988). For example if people are asked to shift their gaze as quickly as possible from one part of an image to another they do not show the scanning effects obtained by Kosslyn and others (Boden, 1988; Pylyshyn, 1984). However, Johnson-Laird has pointed out that the argument can be reversed. Beliefs must also be epiphenomenal because they are 'imagistically penetrable' and so can be interpreted "in a rationally explainable way by images" (Johnson-Laird, 1983). A deeper conflict between the two camps is at the level of functional architecture. The propositional camp (Chomsky, 1980; Fodor, 1975; Fodor, 1987a; Pylyshyn, 1984; Rips, 1986) believes that mind must have an underlying formal representational system that gives rise to *all* high-level thought processes. The imagery camp, though less united, tends to believe in an architecture where even high-level, central processes may have differing, specialized representational systems (Anderson, 1983; Denis, 1991; Johnson-Laird, 1983; Kosslyn, 1980; Marschark & Surian, 1989).

### 1.3.3 Levels of representation

Anderson (1978) attempted to show that the imagery debate cannot be resolved because theories of representation, providing certain assumptions are held, cannot be distinguished from one another empirically (Johnson-Laird, 1980). Thus it may be the case that the differences between theories of representation are not fundamental (Norman & Rumelhart, 1975; Rumelhart & Norman, 1983). Johnson-Laird has taken a different line by suggesting two possible resolutions. The first conclusion is to accept that there may well be some underlying (propositional) representational format. This kind of argument is however not an interesting one for the present thesis because it is not susceptible to empirical investigation by psychologists (Johnson-Laird, 1983; Palmer, 1978). An alternative conclusion is that different representational structures may be used at different *functional* levels of the information processing system (Johnson-Laird, 1983). In keeping with this conclusion Johnson-Laird has proposed three 'high-level' representations which are of interest to psychologists; propositions, mental models and images. Propositions are "mental representations of a verbally expressible proposition", a mental model is a



representation of the structure of a situation and an image corresponds to a “view” of a mental model (the implications of this tripartite distinction are discussed more fully in Chapter 2). It may be the case that these high-level representations are all manifestations of a single primitive representational code. Such a code, however, would still have to explain all the phenomena for which psychologists currently appeal to high-level representations. Whatever the underlying representation may be, it is legitimate to propose high-level representations such as images, mental models and verbal propositions to explain processes like comprehension, planning and reasoning (Anderson, 1983; Johnson-Laird, 1983; Newell, 1990).

This debate has concentrated, to a greater or lesser extent, on explicitly representational issues such as structure. In turn, these have implicated structural and organizational properties of the architectures in which representations are embedded (Anderson, 1983; Pylyshyn, 1989; Rumelhart, 1989). Representational issues are implicit in discussions of cognitive architecture. Perhaps the clearest way of demonstrating this is to focus on the question of process rather than structure or notation. Anderson (1978) has argued that it is in principle impossible to distinguish between different notations for mental representation. Similarly, Palmer (1978) has suggested that fundamental issues of representational structure, and specifically the nature of the isomorphism between the represented world and the representing world may not be open to investigation by cognitive psychologists (he leaves open the question of neuropsychological investigation of this issue). Palmer has also stressed that representational formats must be defined with respect to the actual processing operations that they support.

The use of multiple representational formats in ACT\* marked an apparent reversal of Anderson’s previous position that only a single representational format is necessary (Anderson, 1983). In fact, Anderson was supplementing his previous position by recognizing that form and notation were less important than issues of process and efficiency. Specifically, “different representations are needed ... for different aspects of the same application” (Anderson, 1983). This stance is remarkably similar to the distinction made by Johnson-Laird (1983) between high-level constructs and low-level constructs. Johnson-Laird illustrates his argument with reference to programming structures in a computer. Low-level representations thus correspond to the primitive machine code of the computer. This could just as easily be the propositional representations proposed by Pylyshyn (1984) or the subsymbolic representations of connectionist theorists (Rumelhart, 1989; Smolensky, 1988). High-level constructs such as images or mental models thus correspond to arrays and similar structures in programming languages. For the purposes of the programmer it does not matter that the array is not actually two-dimensional (or that it is implemented in a more ‘primitive’ low-level code), it only matters that it behaves as if it were a two-dimensional structure. In conclusion, it is reasonable to distinguish between multiple forms of representation provided these different forms of representation can be distinguished empirically at some level. For Anderson (1983) they are desirable because some tasks can be carried out more efficiently by a representational system if it has access to special-



purpose representations such as images or temporal strings. Multiple forms of representation are required to cope with the range and complexity of functions carried out by a cognitive architecture (Cohen, 1989; Eysenck & Keane, 1990; Newell, 1990).

## 1.4 The perceptual and structural characteristics of imagery

A central feature of imagery, and for many people the defining feature, is the compelling similarity between the experience of imagery and perceptual experience. As what is imagined becomes more complex the imagery encompasses a greater range (though not necessarily intensity) of sensory, and in many cases, emotional experience. Simple situations often invoke merely visual or auditory imagery (Kosslyn et al., 1978; Paivio, 1986). Tasks like mental paper folding or mental rotation sometimes involve tactile as well as visual imagery (Shepard & Feng, 1972; Shepard & Metzler, 1971). A number of experiments have shown that mentally practising motor skills can lead to improvement in them (Denis, 1991; Paivio, 1986), especially when they are skills with a major cognitive component such as visuo-motor coordination (e.g. catching, hitting or throwing). The common thread in all these examples of imagery is that the experience is, in some way, related to the experience of actually perceiving or reacting to the presence of an object, scene or event.

An important distinction to be made at this point concerns the source of the imagined experience. Modern theories of imagery tend not to differentiate between kinds of images and imagery experience. Some theorists have noted the difference between imagery in episodic and semantic memory (Brewer & Pani, 1983; Conway, 1990). Episodic memory “stores information about temporally dated episodes or events and temporal-spatial relations among these events” (Tulving, 1972). While Tulving (1972) has studied episodic memory in the context of paired-associate learning experiments, his definition is perhaps more appropriate for research into memory for personally meaningful episodes (Brewer & Pani, 1983; Conway, 1990). Semantic memory is “organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, [and] about relations among them” (Tulving, 1972). Most psychologists extend Tulving’s use of semantic memory to cover any abstract knowledge about the world which does not preserve spatio-temporal context (Brewer & Pani, 1983; Hintzman, 1978). Images of specific episodes are more likely to be involuntary memories in the sense described by Ebbinghaus (Brewer & Pani, 1983), and more likely to evoke strong affect. Images from semantic memory are more likely to be actively generated. These differences between episodic and semantic memory images may be related to the distinction (associated with Piaget and Inhelder) between *reproductive* and *anticipatory* imagery. The former are static<sup>1</sup> and derived from memory, while

---

<sup>1</sup> By static Piaget probably means they are not actively manipulated and transformed, rather than that they can not reproduce the experience of a dynamic event or series of events. In either case it is the former



the latter are actively generated and able to be transformed to meet specific goals (Denis, 1991; Kosslyn, 1980).

A major finding in imagery research, that more complex images take longer to generate only applies to images generated from semantic memory (Conway, 1983). Complex images relating to personally meaningful or autobiographical memories are quicker to generate than even simple images from semantic memory. Conway suggests that these autobiographical images are preserved in some kind of "literal" format. In contrast, semantic images are generated, part by part, from a more abstract representational format such as a "propositional list of attributes" (Conway, 1983). These findings are consistent with evidence of imagery use in everyday life. In two diary studies by Kosslyn, Seger, Pani and Hillger (1990) the vast majority of imagery experiences were not directed toward any specific purpose. They concluded that "imagery appears typically to be a spontaneous response to related information". These results are not consistent with the view that imagery experience is only derived from a structural representation (constructed, perhaps, for reasoning or problem solving). The findings suggest that imagery experience is a way of accessing perceptual information about a situation. Representations constructed for reasoning and problem solving do preserve structural and spatial information. Occasionally people report imagery during reasoning, but there is virtually no evidence that visual imagery experience influences reasoning ability. For example, relational problems which are easy to image do not appear to be easier to solve than those which are hard to image (Evans et al., 1993). In one study three-term series problems with concrete face-name associations were significantly harder to solve than those only with name terms (Richardson, 1987).

This interpretation suggests a fundamental distinction between images which have constituent structure and those which do not. This is an important, and neglected, distinction which maps onto the distinction between a mental model and an image. A mental model is an analogical or intrinsic representation which preserves the structural relations between the real-world entities it is representing (Johnson-Laird, 1983; Palmer, 1978). A mental model does not necessarily involve the experience of visual or other imagery. Where imagery does occur, an image corresponds to a view of a mental model (Johnson-Laird, 1983; Marks, 1989). Similarly, imagery can occur in the absence of a mental model and hence without the representation of structural relations,<sup>2</sup> as in the case of vivid autobiographical memories. Conway (1983) has

---

interpretation rather than the latter which is emphasised here. The important point is that similar, but not identical, distinctions between different kinds of images have been proposed before in psychology.

<sup>2</sup> It should be noted that there is a difference between structure that can be derived and structure that is preserved in the representation. Structure can be derived from a purely visual representation such as newspaper



speculated that complex images retrieved from episodic memory may not show the 'classic' imagery effects found for images generated from semantic memory (Kosslyn et al., 1978; Shepard & Metzler, 1971). This distinction may also be related to the separation of visual and spatial components of visual imagery in neuropsychological and neurophysiological studies (Farah, 1988; Farah et al., 1988; Farah, Peronnet, Gonon, & Giard, 1988). Only the spatial component of visual imagery necessitates that structural relations are preserved. In the remainder of this thesis the terms 'image' and 'imagery' are used to describe perceptual and not structural representations.

## 1.5 Memory for objects, scenes and events

In order to understand the nature of mental representation it is essential to consider what demands are placed on the representational system or systems. This requires an understanding not only of traditional laboratory studies of memory, but also of the demand placed on memory in more natural, everyday settings (Baddeley, 1988; Cohen, 1989). Brewer and Pani (1983) have proposed a descriptive or structural account of human memory which tries to take account of the breadth of memorial experience. They conclude that multiple forms of memory representation are required:

" ... consider a typical semantic memory task where a subject is asked "What color is a canary?" and responds correctly. In terms of our analysis the subject's response could have been based on (1) a particularized image, (2) a schema, (3) a generic image, (4) semantic memory, or (5) rote linguistic skill. Clearly, if one is going to construct adequate models of the memory process, one must be sensitive to this issue and attempt to establish what form of representation the subject is using in a given performance ..."

[Brewer and Pani (1983), p. 31]

The purpose of the following sections is not to provide an exhaustive list of different forms or kind of remembering, but to give a flavour of the range of memorial experience.

Early memory research concentrated on the retention of simple entities. In the area of perception, a great deal of research has focused on the recognition of single objects such as dots, lines, or alphanumeric symbols (Bruce & Green, 1985). Even models of categorisation and similarity have tended to deal with collections of discrete, individual objects (Rosch, 1975; Rosch & Mervis, 1975;

---

picture or a computer bitmap. However, nowhere are lines, features or objects represented in the picture. The same is true of a purely visual image or a "literal" representation of an autobiographical episode.



Tversky, 1977). The simplest perceptual models involve the matching of a stimulus pattern (usually standardized by rotation and scaling) to a stored representation of a pattern. These *template matching* models are generally considered too inflexible and too reliant on effective standardization procedures to be useful psychological models, except possibly for very-low level perceptual processes (Bruce & Green, 1985; Palmer, 1978). More sophisticated models make use of *features* rather than *templates*. In the Pandemonium model (Selfridge, 1959) a hierarchy of 'demons' detect progressively more complex features. A third way of representing objects is to use *structural descriptions* which encode the relations between components that make up the object (Bruce & Green, 1985; Palmer, 1978; Reed, Ernst, & Banerji, 1974). The trend from lesser to greater reliance on structure in perceptual processing is clear, but all three methods rely on relatively static representations.

Research on perception and comprehension has suggested that the major problem with representations of objects and scenes is that they are too static. Critics of traditional psychological approaches to perception, such as Gibson (1979) have stressed that the stimulus is a pattern of light intensity over a period of time, not a series of static images. Leading researchers on visual perception have taken static scenes as a departure point for more active representations (Lee, 1980; Marr, 1980; 1982). The notion of schema has also been extended to equally dynamic areas such as coordination and transfer of motor skill (Jordan & Rosenbaum, 1989; Kelso, 1982; Rosenbaum, 1984). Other researchers, have increasingly come to realize that memory is more dynamic than originally conceived (Schank, 1982). For example, research on story recall and comprehension often dealt with relatively static plot units and episodes (Black, 1984; Thorndyke, 1984). Schank's scripts represent typical events like going to a restaurant. Like nearly all schema representations scripts incorporate specific information by the use of default values or routines for calculating and accessing details (Rumelhart & Norman, 1983; Schank & Abelson, 1977). Similarly, research on discourse comprehension has formulated the concept of a situation model or mental model which represents the changing meaning of a text or utterance as it is being understood (Bower & Morrow, 1990; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). These dynamic working memory representations are constructed using real-world knowledge derived from both schematic and episodic knowledge in long-term memory.

The likelihood of a large, probably indefinite, number of 'pre-stored' scripts organized in long-term memory has been questioned both in terms of efficiency and by experimental evidence (Rumelhart, Smolensky, McClelland, & Hinton, 1986; Schank, 1982; Thorndyke, 1984). Because of this Schank has proposed smaller, more abstract script-like entities known as *memory organization packets* (MOPs) which are used to provide expectations from information about earlier events (Rumelhart & Norman, 1983). Instead of DOCTOR and DENTIST scripts there are PROFESSIONAL OFFICE VISIT MOPs. In some (presumably highly familiar) situations scripts will still be used, but in these cases more general information relating to the script will occur in the form of a number of MOPs. Thus scripts are now much more limited in scope and centrality to



cognition (Schank, 1982). What were previously known as scripts by Schank now correspond to *superscripts* ; large structures not stored as a 'chunk' but constructed from high and low level MOPs and related information in memory (Schank, 1982). One advantage of this process of superscript construction is that it is considerably more dynamic and capable of dealing with more complex events in a more flexible way, but with a relatively small number of MOPs. Schank's theory includes MOPs at different levels of generality as well as thematic organization points (TOPs) which capture more abstract properties still (e.g. GOAL PURSUIT) for reminding or high-level analogy (Eysenck & Keane, 1990).

The main observation from a functional analysis of memory and representation is that both working memory and long-term memory have to cope with an enormous range of information. In order to cope with this kind of messy and sometimes unpredictable environment flexible, dynamic systems of representation are necessary:

"We need a dynamic memory system to cope with changing circumstances and a changing physical environment. We need to be able to update the knowledge we have stored and to transform the models of the world we construct in our heads. We have to revise the concepts we have acquired, or throw them out and acquire new ones. Fixed memory structures are liable to become obsolete or to be inappropriate for the current situation. ... Fixed memory structures are uneconomical to store because the same high-level elements need to be reduplicated in many representations. Dynamic memories are readily revised, updated, and modified, whereas fixed memories would rapidly become redundant in a changing world."

[Cohen (1989), p. 218-219]

The sheer amount of information available in the environment requires that memory preserve or derive structure not only within a given individual memorial event, but in terms of the way memorial events are organized in long-term memory. It is tempting to conclude that dynamic representations are only constructed and manipulated in the temporary, fluid medium of primary or working memory. However, long-term memory must reflect not only the static end-products of, but also, at the very least, the intermediate stages of working memory processes. Long-term memory also has to reflect the changes in working memory representations over time and from situation to situation. There is ample evidence in areas such as navigation or language understanding that long-term memory is able to integrate information acquired over an extended period of time or from different situations (Gernsbacher, 1991; Levine et al., 1982; Thorndyke & Hayes-Roth, 1982; van Dijk & Kintsch, 1983). Both working memory and long-term memory have to reflect the demands of a complex, dynamic, environment.

## 1.6 The influence of working memory processes on long-term memory

One of the most influential ideas in contemporary memory research is that encoding processes acting in working memory determine the durability of memory traces in long-term memory. Any investigation of representation in long-term memory would therefore be incomplete without some discussion of how encoding processes influence recognition or recall. A straight-forward discussion of encoding, however, is problematic. This is because tests of memory for encoding processes are not independent of conditions at retrieval. The next two sections discuss how performance on a memory test is a product of encoding processes, retrieval cues or the interaction of the two.

### 1.6.1 Memory and encoding processes

In 1972 Fergus Craik and Robert Lockhart proposed that the durability of memory traces was primarily the result of progressively 'deeper' and more 'meaningful' processing during encoding. This levels or depth of processing account was able to account for experiments that showed that perceptual processing (e.g. deciding whether a word started with the letter 'T') produced worse retention than semantic processing (e.g. deciding whether a word described an animate or inanimate object). This processing account of memorability proved to be problematic; the chief criticism being that it is not always possible to define 'deep' or 'semantic' processing *a priori*. Some experiments required *post hoc* judgements to explain results, others demonstrated that perceptual tasks could result in superior recall to apparently semantic tasks (Bransford et al., 1979; Craik, 1979; Eysenck, 1979; Horton & Mills, 1984; Jacoby & Craik, 1979; Morris et al., 1977). Rather than being abandoned altogether the levels-of-processing framework has been expanded and revised to incorporate a number of different influences of processing at encoding and retrieval (Craik, 1979; Jacoby & Craik, 1979). Depth of processing refers to qualitative changes in a memory trace brought about by processing and still remains difficult to define and hence test empirically, though it retains its intuitive or metaphorical appeal (Craik, 1979; Craik & Lockhart, 1972; Eysenck & Keane, 1990; Horton & Mills, 1984).

Several different kinds of processing accounts have emerged from work within, or in response to, the levels-of-processing framework (Craik, 1979; Horton & Mills, 1984; Smyth, Morris, Levy, & Ellis, 1988). Some accounts which stress the *amount* of processing which an input receives. For example, if people are asked to classify a list of words on two dimensions (e.g. whether something is consumable and whether something is a solid) words with both properties are remembered best, while words with neither are remembered worst (Johnson-Laird, Gibbs, & de Mowbray, 1978). Other accounts stress the *type* of processing involved. Good examples of these are relational and item-specific processing proposed by Marschark and others (Hunt & Einstein, 1980; Marschark, 1985; Marschark & Surian, 1989). Both these accounts are subsumed



under *elaboration* accounts which propose it is the combination of complexity and amount of processing which underlie levels-of-processing effects (Anderson & Reder, 1979; Eysenck, 1979; Reder, 1979; Reder, 1980). According to this view depth of processing effects occur “by changing the number and type of elaborations stored” (Anderson & Reder, 1979). A related, but slightly different account, is based on the *effort* expended during encoding (Horton & Mills, 1984). Effort is usually defined as the amount of processing capacity brought to bear on a task. Using divided attention procedures it is possible to show that effort is a better predictor of retention than time spent on the task (Eysenck & Eysenck, 1979).

*Distinctiveness* accounts propose that certain kinds of processing can lead to memory traces which are more easily differentiated from other memory traces and hence more easily remembered (Eysenck, 1979; Horton & Mills, 1984). *Congruity* accounts (see Craik, 1979) stress the importance of the overlap between encoding and retrieval operations. One of the most striking demonstrations of this was made by Morris, Bransford and Franks (1977). They compared a memory for words learned on a ‘deep’ semantic task (judging the whether the noun formed part of a meaningful sentence) or a ‘shallow’ perceptual task (making a rhyming judgement). Memory for semantically processed words was superior for a standard recognition task, but significantly poorer for a rhyming recognition task. These results were explained in terms of transfer-appropriate processing; people remember the activity or processing they engage in not just the input itself (Morris et al., 1977). Both distinctiveness and congruity place enormous emphasis on the relationship between encoding and retrieval conditions. In the case of distinctiveness an input cannot be classed as distinctive in isolation but only within a given context (Horton & Mills, 1984). In addition, distinctiveness will also be influenced by the type and amount of processing during encoding because some kinds of processing can give rise to more distinctive memory traces than others (Craik, 1979; Kulhavy, Stock, Woodard, & Haygood, 1993). It is almost certain that different kinds and amounts of processing during encoding and retrieval give rise to apparent ‘depth’ of processing effects.

### 1.6.2 The influence of retrieval cues on recall

In discussing how encoding processes influence long-term memory it should be apparent that in every case what is encoded must also be retrieved or in some cases reconstructed. For this reason many of the processes that influence encoding long-term memory are also implicated at retrieval. This position is stated in its strongest form in the *encoding specificity principle* (Tulving & Thomson, 1973). According to the encoding specificity principle, if information cues recall then that information must have been encoded onto the memory trace during the original learning episode. Stated in this form the principle is circular; any cue that promotes recall or recognition must *de facto* have been encoded (Eysenck & Keane, 1990). Rather than focus on the “informational overlap” (Tulving, 1979) between encoding and retrieval, because many of these



factors have been raised in the discussion of encoding processes, the remainder of this section describes how different cues promote or fail to promote retrieval. Looking at the effectiveness of different kinds of cues demonstrates how the retrieval context can have a significant effect on recall. It will also suggest that a simple acceptance of the encoding specificity principle is unwarranted.

At this point it is helpful to distinguish between two traditions of human memory research. In one tradition people learn lists of verbal stimuli (e.g. words or nonsense syllables). At retrieval people are presented with cues such as the name of the list or paired associate learned with the list. In this verbal learning tradition the influence of retrieval context on recall is usually minimal; what is important is the overlap between the retrieval context and the encoding context (Craik, 1979; Jones, 1982; Tulving, 1979). In the second tradition people are presented with more complex stimuli such as stories in what are now termed more *ecologically valid* situations (Bartlett, 1932; Brewer & Pani, 1983; Cohen, 1989). In these experiments retrieval context has a marked effect on recall; it is argued that people reconstruct the original learning material by using the retrieval context to generate general knowledge in the form of schema from long-term memory (Anderson & Pichert, 1978; Bartlett, 1932; Cohen, 1989). However, it would be misleading to suggest that these two traditions are reflected in entirely separate, distinct or opposed fields of research. One example of the merging of the two approaches is in the *generate-recognition* theory of recall. This theory can be traced back at least as far as William James (Brown & McNeill, 1966; James, 1890; Watkins & Gardiner, 1979). According to generate-recognition theory when presented with a memory cue people attempted to directly match that cue with an element in long-term memory. In the simplest case (recognition) the cue and retrieval target are identical and match. In the likely event that the cue and target are not identical, people use the cue (e.g. "BARK") to generate or search for additional cues (e.g. "TREE" or "DOG") until the correct target is recognized (e.g. "DOG" is matched with "DOG"). Versions of this theory account well for the ways people actively search and make use of their memory, but have problems, for example, in explaining recognition failure and context effects in recognition memory (Jones, 1978; Watkins & Gardiner, 1979). Recognition failure is said to occur when people are able to recall something but not recognize it (Tulving & Thomson, 1973; Watkins & Gardiner, 1979).

One of the most successful recent theories of recall provides an account which attempts to reconcile models which propose recall is based on the degree of "informational overlap" between cue and target (Tulving, 1979; Tulving & Thomson, 1973) and those where retrieval is always preceded by recognition (Jones, 1978; Watkins & Gardiner, 1979). This dual mechanism theory (Jones, 1978; 1982; 1987) proposes that there are two routes to recall differentiated by two kinds of retrieval cue. The first kind of retrieval cue contains information encoded during the original learning episode and is termed an *intrinsic* cue. The second kind of retrieval cue contains no information from the original learning episode and is termed an *extrinsic* cue. Intrinsic cues thus



provide direct access to the memory trace. Extrinsic cues provide indirect access in the manner of generate-recognition theory discussed earlier. The dual mechanism theory has been successfully used to predict the relationship between levels of recognition and recall in many data sets from paired associate learning experiments (Jones, 1978; 1987; Le Voi, 1984). It has also been argued that it is the only theory which can be used to derive algebraically the Tulving-Wiseman equation for the relationship between the recognition and recall (Jones, 1987). In addition, the prediction that the availability of the indirect, extrinsic route increases recall has also been supported (Jones, 1982). Problems with the model have centred not so much on the availability of the two routes, but rather on the relationship between them, and in particular whether they are stochastically independent or exclusive (Jones, 1982; 1987).

## 1.7 Conclusions

This chapter has attempted to address the notion of representation, and in particular the representation of spatial information, in human memory. The approach that has been adopted is a functional one, based primarily on the role or roles which representation plays in information processing theories. Firstly, it has been argued that human memory may require more than one system of representation. These forms or systems of representation can be distinguished by traditional cognitive psychological methods and are defined by the kinds of special-purpose, high-level procedures able to operate on them rather than in terms of a primitive representational code. Secondly, it has been argued that memory must be flexible and dynamic in order to cope with the range and complexity of demands placed on people in everyday situations. Thirdly, it has been argued that long-term memory traces reflect the processes that construct and manipulate representations in working memory. The memory traces formed by these processes interact with cues available in the retrieval context when memory performance is tested. The common theme running through all these issues is the relationship between special-purpose representations in working memory and memory traces encoded in long-term memory. Of particular interest is the notion of a spatial mental model. It has already been noted that a spatial mental model is a representation of the spatial structure of a situation and can be distinguished from an image (a representation of the perceptual characteristics of a situation) or a linguistic proposition (a representation of a verbal description). In the next chapter theoretical and experimental evidence in relation to spatial and non-spatial mental models will be discussed in more detail.

## Chapter 2 - The representation of spatial mental models in memory

This first part of this chapter is concerned with the representation of mental models in working memory. The purpose of this discussion is to present evidence for and against the thesis that mental models are also represented in long-term memory. While the theory of mental models advanced by Johnson-Laird (1980; 1983) is a general theory of mental representation the evidence will, where possible, focus on mental models of spatial situations (spatial mental models). Experimental and theoretical work on mental models of non-spatial situations is also discussed where it is important to the theory or where it complements available evidence on the representation of spatial mental models in working memory or long-term memory. The chapter ends with an outline of the processes that constitute mental modeling. This outline is motivated by the theoretical and empirical discussion presented earlier in the chapter and is intended to predict how mental modeling processes in working memory should influence long-term memory.

The fundamental premise of mental models theories is that thinking and understanding depend on the ability of the mind to construct 'working models' of real world phenomena. According to Johnson-Laird (1983) the first modern formulation of this thesis is to be found in *The Nature of Explanation* (Craig, 1943). The importance of Craig's contribution is that he considered human beings as processors of information. He proposed that the mind constructs a model which "has a similar relation-structure to that of the processes it imitates." Though the digital computer had not yet been invented, Craig's language is very similar to that of later psychologists working within a computational or cognitive framework. Recent formulations of the mental model thesis in psychology have focused on the kind of mental representation necessary or desirable either in reasoning or language understanding. Of particular importance is the work of Johnson-Laird (1980; 1981; 1983; 1989; 1993) because it draws together arguments and evidence from both these areas of research.



## 2.1 The theory of mental models

The theory of mental models (Johnson-Laird, 1983) is concerned with uniting three areas of cognitive science research; the study of inference, language understanding and consciousness. Johnson-Laird argues that the notion of a mental model is fundamental to psychological research in both language comprehension and reasoning. Furthermore, the ability to construct mental models and to embed them within other mental models may be an essential component in self-awareness (Johnson-Laird, 1983; 1988a; 1988b). A full treatment of all these ideas is beyond the scope of this discussion. However, an overview of the nature of mental models and their role in reasoning and comprehension will be provided.

### 2.1.1 What is a mental model?

According to Johnson-Laird a mental model is not simply any representation that models real world situations, but a representation that models those situations in a particular way. Johnson-Laird (1989) offers the following working definition:

“ ... a mental model can be defined as a representation of a body of knowledge - either long-term or short-term that meets the following conditions:

1. Its structure corresponds to the structure of the situation it represents.
2. It can consist of elements corresponding only to perceptual entities, in which case it may be realized as an image, perceptual or imaginary. Alternatively it can contain elements corresponding to abstract notions; their significance depends crucially on the procedures for manipulating models.
3. Unlike other proposed forms of representation, it does not contain variables. Thus a *linguistic* representation of, say, *All artists are beekeepers* might take the form

For any x, if x is an artist, then x is a beekeeper.

In place of a variable, such as ‘x’ in this expression, a model employs tokens representing a set of individuals.”

[Johnson-Laird (1989) p. 488]

Mental models are contrasted with propositional representations and images. Johnson-Laird’s use of propositions is confined to linguistic representations close to the structure of written or spoken discourse. There is at present a large and growing body of evidence to distinguish between the representations that are structurally similar to the *situations* they represent and representations that are similar to the *language* of the text or discourse from which they are

constructed (Bower & Morrow, 1990; Bransford & Franks, 1971; Glenberg et al., 1987; Johnson-Laird, 1981; Perrig & Kintsch, 1985; Tversky, 1991; van Dijk & Kintsch, 1983). Many of the experiments take advantage of the inability of mental models to directly represent indeterminacy. For example, in order to represent the possible states of affairs described by the relation 'next to' mental models would have to be constructed for each plausible alternative (e.g. 'to the right of', 'to the left of', 'above' and so on). By contrast, a propositional representation can encode this relation without explicitly representing two or more alternatives. Images correspond to views of mental models. Mental models may be accompanied by visual (or for that matter, other kinds of) imagery experience. Imagery, though, is not a necessary requirement - in fact, as the third point in the above quotation suggests, abstract or conceptual models may be difficult (or even impossible) to visualize. Unlike the distinction between propositions and mental models the relationship between mental models and imagery is left vague in the theory. One plausible interpretation of this relationship is that imagery corresponds to the generation or retrieval of perceptual information about a situation. Imagery would thus play an important role in 'consulting' a spatial mental model where visual information such as colour, perspective or texture are required for to meet a goal or solve a problem. Johnson-Laird's theory of mental models made an important contribution to the 'imagery debate' and attempted to clarify the notion of mental representation in cognitive science (refer to Chapter 1 for a more detailed discussion of these points).

## 2.1.2 Mental models and reasoning

Mental models can be constructed for a number of different purposes. Johnson-Laird concentrates on three main areas of cognition where the notion of mental models is particularly attractive. The first area is that of reasoning. Traditional approaches to reasoning in psychology and philosophy have (implicitly or explicitly) focused on what Johnson-Laird has termed the 'doctrine of mental logic' (Johnson-Laird, 1983; 1986; Johnson-Laird & Byrne, 1991). Proponents of mental logic assert that human thinking is fundamentally a manifestation of an innate or acquired logic or logics; a set of rules and a set of procedures for applying them (Braine, 1978; Henle, 1962; Piaget & Inhelder, 1958; Rips, 1983). Psychological theories of reasoning based on logic face difficulties attempting to account for logical errors, biases, content effects and so on. Typically they also involve non-logical mechanisms for extracting the premises of an argument and for generating an output from the conclusions provided by the logical, inferential processes. The mental models theory proposes instead that people reason using representations of the situations described by the premises. Johnson-Laird and his colleagues have shown that semantic procedures rather than logical, syntactic ones can be used to reason deductively (Johnson-Laird, 1983; 1991; 1993). Johnson-Laird suggests that people attempt to construct a series of models of possible conclusions in working memory. Many deductive errors occur when people fail to search for counterexamples or because of working memory limitations (Johnson-Laird & Byrne, 1991; Oakhill & Johnson-Laird, 1985b). Similarly, belief bias and content effects can arise



through the construction and consultation of models influenced by real world knowledge (Oakhill & Garnham, 1993; Oakhill, Garnham, & Johnson-Laird, 1990; Oakhill & Johnson-Laird, 1985a; Oakhill, Johnson-Laird, & Garnham, 1989). A considerable body of evidence now supports the claim that model-based reasoning provides a more plausible account of reasoning competence, biases and errors than existing logic-based accounts (Baron, 1988; Boden, 1988; Cohen, 1983; de Vooght & Vandierendonck, 1993; Evans, 1991; Evans et al., 1993; Eysenck & Keane, 1990; Galotti, 1989; Gilhooly, 1988; Manktelow & Over, 1992; Newell, 1990; Newstead & Evans, 1993; Ormerod, Manktelow, & Jones, 1993). A detailed account of how people construct mental models to reason with spatial descriptions is provided later in this chapter (see section 2.1.7).

### 2.1.3 Mental models and discourse comprehension

While mental models have undoubtedly had their biggest impact in the area of reasoning they have also made a key contribution to the study of discourse comprehension. The mental model theory proposes that people construct not only a mental representation of the language of a text (linguistic propositions) but also a representation of the situation described by the text. This distinction built on the classic work of early construction-integration theorists (Bransford et al., 1972; Bransford & Franks, 1971). Mental models are necessary to explain superior memory for gist, and causal or other inferences made about the situation described by the discourse (Bower & Morrow, 1990; Garnham & Oakhill, 1992; van Dijk & Kintsch, 1983). Propositional representations are a necessary stage in the construction of a mental model and are needed to explain verbatim memory and memory for poorly comprehended discourse. Experimental work by Johnson-Laird and his colleagues has provided considerable support for this distinction (Ehrlich & Johnson-Laird, 1982; Garnham, 1981; 1987; Johnson-Laird, 1981; Johnson-Laird, 1983; Oakhill & Johnson-Laird, 1984). Johnson-Laird's theory of mental models also fits well with work by other researchers who also argue that comprehension often requires representations of situations rather than representations of the text (Bower & Morrow, 1990; Gernsbacher, 1991; Schank, 1975; van Dijk & Kintsch, 1983). An understanding of discourse comprehension is important to this thesis for two reasons. Firstly, information about spatial situations is often derived from discourse. Secondly, it is likely that many cognitive processes involved in discourse comprehension are general cognitive processes (Gernsbacher, 1990; 1991).

The theory of discourse comprehension proposed by van Dijk and Kintsch differs from that of Johnson-Laird in several important respects. It offers a detailed account of discourse comprehension rather than a general account of how people construct and transform representations of the world. A great deal of the theory and the subsequent research inspired by it addresses how people construct representations of the text rather than representations of the



situations described by the text (Kintsch, 1988; van Dijk & Kintsch, 1983). Kintsch (1986) distinguishes between three separate (though not wholly independent) levels of analysis:

“(a) the processes concerned with the parsing of the text; (b) the establishment of a coherent representation of the meaning of a text, both at the local and global levels; and (c) the integration of the text content into the comprehender’s knowledge system.”

[Kintsch (1986), p. 89]

Each level of analysis corresponds to a particular representation of text; the *surface* form, the *textbase* and the *situation model*.

The surface form of the text tends to be short-lived, but is essential for the formation of the textbase. The textbase corresponds closely to the linguistic propositions representation proposed by Johnson-Laird and Garnham (Garnham, 1987; Johnson-Laird, 1983). One key distinction is that for Johnson-Laird the propositional representation is closer to the surface form of the text than the textbase proposed by van Dijk and Kintsch. According to their account the textbase may include higher level structures (e.g. macropropositions created by combining linked propositions) in addition to individual propositions resembling sentences or phrases from the original text. However, the two theories have more similarities than differences. For example, the van Dijk and Kintsch (1983) model is able to account for both views by appealing to the notion of processing cycles during discourse comprehension. Early in processing the textbase may retain many features of the surface form, later in processing the textbase may be sufficiently elaborated to include representations of the macrostructure as well as the microstructure of the text. A second point of departure from the mental models theory of Johnson-Laird and his associates lies in the concept of the situation model. Unlike Johnson-Laird, van Dijk and Kintsch have not attempted to address the representational status of the situation model directly. One possibility is that though the situation model does not explicitly represent the language of a text, its underlying representation is propositional (Kintsch, 1988; van Dijk & Kintsch, 1983). Elsewhere, Kintsch raises the possibility that a situation model may also take the form of an image or a temporal string (Anderson, 1983; Kintsch, 1986; 1988; 1994).

## 2.1.4 The characteristics of mental models

### Intrinsic representation of structure

As already noted, the structure of relations within a mental model corresponds to the structure of the relationships in the situation being modeled. This is the defining feature of a mental model and is inspired by the work of Kenneth Craik (1943). It is very similar to the notion of intrinsic



representation proposed by Palmer (1978). As will be shown, this property gives rise to several of the other important properties of mental models.

### Working memory limitations

The construction and manipulation of a mental model is limited by the capacity of working memory.<sup>1</sup> This characteristic arises both from empirical research (Baddeley, 1986; Daneman & Carpenter, 1980) and from *a priori* arguments that any representational system must have some limit on processing power (Johnson-Laird, 1983).

### Explicit representation of structure

All relations intrinsically represented in a mental model are by definition explicit. A mental model cannot directly represent indeterminacy. Indeterminacy can be represented indirectly (e.g. by explicitly representing two or more possible alternative situations). For this reason an explicitly represented situation requires as great or greater cognitive resources than an implicitly represented situation. Explicit representation is advantageous when the goals of the representational system require it to use knowledge that is currently implicit (e.g. a deduction from a set of premises). Where the knowledge held by a representational system contains indeterminate relationships explicit representation becomes inefficient (Johnson-Laird, 1983; Newell, 1990). A more efficient approach might be to abandon mental model construction and rely on an intrinsic representation such as a set of linguistic propositions.

### Incompleteness

The explicit nature of relations represented in a mental model makes the representation of indeterminacy difficult, or because of capacity limitations, impossible. Similarly, many situations represented by a system are potentially too complex to be represented by a limited capacity system. One solution to this problem is to omit relations from a mental model; a mental model will usually contain only the minimum number of relations necessary to represent the relevant characteristics of a situation. It follows that it is not a simple matter to predict or otherwise specify the contents of a mental model held by an individual. The nature of the incompleteness of a given mental model will be influenced by the processes of selection or abstraction employed by an individual.

---

<sup>1</sup> Working memory is used here in its general sense rather than as a reference to a specific theory of working memory.

### Abstract representation of structure

Mental models are abstract in the sense that they are not tied to a specific modality. Unlike images, mental models are not restricted to a particular viewpoint or perspective (Johnson-Laird, 1983; Tversky, 1991). According to Johnson-Laird mental models are capable of representing abstract relations such as negation and disjunction<sup>2</sup> (Johnson-Laird & Byrne, 1991), although as has already been argued, a mental model is explicit and therefore not capable of representing highly abstract properties such as indeterminacy (also see section 2.3.5). A mental model is thus more abstract than an image but less abstract than a proposition (Marks, 1990).

### Dynamic representation of structure

Mental models are constructed and updated in sequence. When an alternative mental model is constructed during reasoning the existing mental model is usually abandoned and a new model constructed (Johnson-Laird & Byrne, 1991; Oakhill & Johnson-Laird, 1985b). When a situation being represented changes (e.g. during discourse comprehension) the current mental model is updated to reflect those changes (Bower & Morrow, 1990; Glenberg et al., 1987). Mental models are transient, dynamic entities. The term *mental model* often denotes a 'snapshot' of the processes involved in constructing, maintaining and updating a mental model. The dynamic characteristic of mental models helps distinguish them from the relatively static way imagistic representations are portrayed in the literature, as well as being desirable in order to capture the dynamic properties of everyday memory (Cohen, 1989).

### Cognitive economy

It has already been argued that explicit representations such as mental models are less efficient than propositions at representing certain kinds of information. However, Alan Newell has claimed that mental models are easier to process than propositional representations (Newell, 1990). The apparent contradiction arises from the use of two different levels of analysis. The claim that mental models are less efficient than propositions at representing indeterminacy is a claim about the representation of a specific situation. Newell is making a claim about the processing power of a representational system (SOAR). Mental models in SOAR are powerful because they are incomplete; only relations explicitly represented in the representation need to be updated in the light of other changes in its working memory:

---

<sup>2</sup> It should be noted that some psychologists disagree that a mental model can represent negation or disjunction (Holyoak, 1993). However, this thesis does not require that negation or disjunction are represented in spatial mental models.



“Model representations are *limited* in scope but *easy* to process. Simplicity is obtained because models adhere to a *structure-correspondence principle*. ... No completeness or exhaustiveness is implied.”

[Newell (1990) p. 390]

Newell argues that, unlike mental models, propositional representations suffer from what is known as the ‘frame problem’ (Fodor, 1983; Fodor, 1987b; Newell, 1990). The frame problem concerns how to decide which propositions to update in a large representational system when one or more represented states change. Mental models avoid the frame problem at the cost of incompleteness (or ‘scope’ in Newell’s terminology).

### **Mental models support consultation processes**

Mental models explicitly represent the structure of the situations they represent. They do this at the cost of completeness or scope. In return, inexpensive “match-like and counting processes suffice for working within the model” (Newell, 1990). These processes all depend on some form of ‘consultation’ process. The utility of mental models depends to a high degree on these consultation processes. Deduction in a mental model requires that an appropriate mental model is constructed and that novel explicit information contained in the model is ‘read off’ (Johnson-Laird & Byrne, 1991). Similarly in non-deductive reasoning and problem solving an appropriate mental model may support ‘insight’ (Eysenck & Keane, 1990; Montgomery, 1988). Glenberg and colleagues have proposed an mental model consultation process called ‘noticing’ (Glenberg & Langston, 1992). When a mental model is updated the relationship between the updated element and those in proximity to it is noticed. The noticed relationship is recorded in long-term memory, facilitating the retention of important or useful information.

### **Accessibility to consciousness**

Johnson-Laird (1983; 1988a) has argued that mental models may play an important role in consciousness. Elsewhere (Johnson-Laird & Byrne, 1991) it has been suggested that the “... tokens of mental models ... may not be directly accessible to consciousness.” The accessibility of mental models to consciousness is therefore in some doubt. It is proposed here that the content of any mental model is in principle accessible to consciousness. It follows that conscious recollection and verbal report may provide evidence for the processes involved in the construction and manipulation of a mental model. It is assumed, however, that accessibility is limited by the transient, dynamic properties of the mental model or by working memory limitations. Proponents of the mental models theory argue that errors in reasoning reflect a subset of the possible models of the premises (Johnson-Laird, Byrne, & Schaeken, 1992). This line of argument suggests that the content of a mental models is, in principle, accessible to consciousness (Evans, 1990; Evans, 1993).

### **2.1.5 Alternative conceptions of mental models**

A separate strand of mental model research is concerned with the theories and models people generate when interacting with or attempting to understand complex devices or systems. The existence of alternative conceptions of mental models is a source of confusion, particularly as there are similarities between the different strands of research. The term “mental model” has been used in several different (though not necessarily contradictory) ways (Brewer, 1987; Johnson-Laird, 1989). One approach treats mental models as complex knowledge structures corresponding to complex domain (or other real-world) knowledge which people hold (Gentner & Stevens, 1983). This approach is popular in human-computer interaction and other research on understanding complex systems. Often these models are characterized as particularly rich schema-like knowledge structures, though many researchers prefer to distinguish these mental models from scripts or frames by stressing their dynamic, transient nature and the ability of subjects to explore problem areas by “running” these models (Holland, Holyoak, Nisbett, & Thagard, 1986; Norman, 1983; Payne, 1988; Payne, 1991; Rumelhart & Norman, 1983).

These formulations of a mental model differ to a greater or lesser extent from that proposed by Johnson-Laird. Johnson-Laird's mental models are high-level, analog structures in working memory. Definitions of different kinds of ‘mental models’ thus lie at two ends of a single dimension. At one end lie relatively static long-term knowledge structures (Minsky, 1975; Schank & Abelson, 1977). At the other end are high-level, transient, dynamic, capacity-limited knowledge structures in working memory (Baddeley, 1992; Garnham, 1987; Johnson-Laird, 1983; Kosslyn, 1980; van Dijk & Kintsch, 1983). Somewhere between these falls the use of mental models to describe the conceptual understanding of real-world situations or naïve physics (Anderson, Tolmie, Howe, Mayes, & Mackenzie, 1992; Gentner & Collins, 1990; Gentner & Gentner, 1983; Hayes, 1979; Hayes, 1985).

### **2.1.6 Mental models as representations of space**

Johnson-Laird (1983) distinguishes between physical and conceptual mental models. The distinction is useful but also a little misleading. The structure and characteristics of a mental model vary according to the structural relations of the situation it is modeling. This in turn will be influenced by the way an individual conceptualizes a given domain. Many people try to conceptualize abstract problems in physical terms, for example, through the aid of metaphors or analogies. Of course the particular analogy adopted may prove either fruitful or misleading depending on the nature of the task (Gentner & Gentner, 1983).



Conceptual mental models are classified as monadic, relational, meta-linguistic or set-theoretic (Johnson-Laird, 1983). Conceptual models are required for comprehending or reasoning with abstract relations. While some abstract relations, such as disjunction, are relatively infrequent, others, such as negation, are very common. *Monadic* models represent identity or non-identity relations among finite sets of individuals. *Relational* models represent finite sets of mappings between tokens of one set and another. *Meta-linguistic* models represent abstract relations between abstract linguistic expressions and elements within a mental model of any type. *Set-theoretic* models permit tokens directly representing sets rather than individuals.

Physical mental models are classified as relational, spatial, temporal, kinematic, or dynamic (Johnson-Laird, 1983).<sup>3</sup> This typology is ordered approximately from low to high complexity. A *relational* model is a finite, static set of tokens and relations (or 'frame') representing physical entities in the world. A *spatial* model is a relational model containing only spatial relations (in two or three dimensions). A *temporal* model consists of a series of spatial 'frames' representing the temporal order of a sequence of events. A *kinematic* model is a temporal model that is psychologically continuous and capable of representing changes and movements in real time. A *dynamic* model is a kinematic model where some or all of the represented relations are causal.

In this thesis the term *spatial mental model* is taken to mean any physical mental model which includes spatial relations between tokens. It is, both in principle and in practice, difficult to separate the spatial, causal and temporal properties of many spatial situations. For example a route is primarily considered a spatial representation, yet is inevitably both temporally and causally structured. The processes of constructing and updating a mental model are sequential. It is possible that more complex physical mental models take advantage of the sequential nature of these processes in order to represent causality and temporal order. The typology of physical mental models described here reflects differences of degree or complexity rather than differences in kind. The mental model account of temporal reasoning is closely based on the mental model account of spatial reasoning (Schaeken, Johnson-Laird, & d'Ydewalle, 1993). Other researchers have proposed and found evidence for a 'spatialized' conception of time in reasoning with mental models (de Vooght & Vandierendonck, 1993; Vandierendonck & de Vooght, 1992).

---

<sup>3</sup> Johnson-Laird (1983) also includes images as types of physical mental model. However, as an image is described as a view or projection from an underlying spatial or kinematic mental model it has not been included as a separate classification here.

2.1.7 Mental models and spatial reasoning

A summary of the mental model account of spatial reasoning is provided here to demonstrate the process of reasoning with mental models and to provide empirical evidence of the representation of spatial mental models in working memory. The account is derived from work on two-dimensional spatial reasoning by Byrne and Johnson-Laird (1989; 1991). The main challengers to mental model accounts of spatial reasoning are logic or rule-based accounts such as those of Hagert (1984).

In spatial reasoning tasks people are given a number of premises describing spatial relationships between two items, for example 'A is to the right of B' or 'C is in front of D'. People engaged in the task are asked to determine what follows from these sets of premises.

<u>Single Model Valid Conclusion Problem</u>	<u>Corresponding Model:</u>
A is on the right of B	
C is on the left of B	C    B    A
D is in front of C	D    E
E is in front of B	
What is the relation between D and E?	
 <u>Multiple Model Valid Conclusion Problem</u>	 <u>Corresponding Models:</u>
B is on the right of A	i)    C    A    B
C is on the left of B	D       E
D is in front of C	
E is in front of B	ii)   A    C    B
What is the relation between D and E?	D    E
 <u>No Valid Conclusion Problem</u>	 <u>Corresponding Models:</u>
B is on the right of A	i)    C    A    B
C is on the left of B	D    E
D is in front of C	
E is in front of A	ii)   A    C    B
What is the relation between D and E?	E    D

Figure 2.1: Spatial mental models constructed from the premises of three types of problem.



Byrne and Johnson-Laird argue that people solve these problems by constructing a mental model of the spatial relations between the items in the premises. First, each premise is read and the meaning of the relation understood. Second, tokens corresponding to items in the premise are entered into the mental model consistent with the meaning of the relationship given in the premise. This continues until all premises have been read and the spatial mental model is complete. At this point a putative conclusion is drawn. However, not all two-dimensional problems can be solved by the construction of a single spatial mental model. In order to be certain of reaching a valid conclusion (or to be certain that there is no valid conclusion) a reasoner should search for counterexamples by attempting to construct different spatial mental models derived from the same premises. Figure 2.1 gives examples of models derived from single model, multiple model and no valid conclusion problems.

On the basis of this account Byrne and Johnson-Laird (1989) predicted that single model problems with a valid conclusion should be easier than multiple model problems with a valid conclusion. In addition, problems with valid conclusions should be easier than problems with no valid conclusion (because all the models of the problem have to be considered to appreciate that nothing follows). Although rule-based accounts also predict that problems with no valid conclusions are the most difficult (because all potential derivations need to be attempted before reaching a conclusion) rule theories predict no difference between single and multiple model problems. These predictions were borne out by the two experiments reported by Byrne and Johnson-Laird (1989). Their second experiment was particularly interesting because the valid conclusion problems they used were chosen to give opposite predictions between the mental model and the rule theory. The single model problems were chosen because they have long derivations (in terms of the number of rules that need to be applied) according to the rule theory. The multiple model problems were chosen because they had short derivations. The rule theory predicts that problem difficulty should increase with the number of rules that need to be applied. However, Byrne and Johnson-Laird found that the length of the derivation required to solve a valid conclusion problem did not predict difficulty.

One of the most common defences of rule-based accounts of reasoning is that different predictions about the difficulty of problems are obtained by postulating different sets of deduction rules. Johnson-Laird and Byrne (1991) have detailed an elegant counterargument to this. They demonstrate that a rule theory can be constructed that makes the same predictions as the mental model theory in their second experiment (Byrne & Johnson-Laird, 1989), but it would fail to explain performance on the problems in their first experiment (Johnson-Laird, 1993). Hence, they argue that no rule theory can account for the overall pattern of difficulty observed. There will always be two-dimensional spatial reasoning problems where the rule theory

departs from the predictions derived from the mental model account and in those cases the experimental evidence supports the mental models account.

## 2.2 Mental models and theories of working memory

Work by Alan Baddeley and colleagues (1992; 1986; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993) provides a useful framework within which to consider the relationship between mental models and working memory. Recent experimental studies have begun to investigate this relationship in more detail (de Vooght & Vandierendonck, 1993; Gilhooly, Logie, Wetherick, & Wynn, 1993; Morra, 1989; Morra, Pascual-Leone, Johnson, & Baillargeon, 1991; Toms, Morris, & Ward, 1993).

### 2.2.1 A specific model of working memory

William James (1890) distinguished between primary memory (corresponding to the immediate content of consciousness) and secondary memory (equivalent to short- and long-term memory stores). Cognitive psychologists have proposed a number of theories which postulate a working memory which contains only temporary representations relevant to current conscious cognitive activity. The specific model of working memory proposed by Baddeley and Hitch (1974) made a tripartite distinction between a *central executive* (a multi-modal, limited capacity processor resembling attention), a *phonological or articulatory loop* (a temporary store of speech-based information) and a *visuo-spatial sketch pad* for generating and manipulating visual or spatial information. It is argued that each of these components has a finite, limited capacity.

The central executive serves as a central working memory resource and acts as an executive or regulatory system over slave systems such as the phonological loop and visuo-spatial sketchpad:

“Some of its primary functions are regulatory in nature: It coordinates activity within working memory and controls the transmission of information between other parts of the cognitive system. In addition, the executive allocates inputs to the phonological loop and sketchpad slave systems, and also retrieves information from long-term memory.”

[Gathercole and Baddeley (1993), p.5]

The central executive is thought to be involved in tasks such as reasoning, problem solving, mental arithmetic, random letter generation, reading, language comprehension, planning and recollection (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Baddeley, 1986; Baddeley &



Wilson, 1988; Daneman & Carpenter, 1980; Gathercole & Baddeley, 1993; Just & Carpenter, 1992; Morra, Moizo, & Scopesi, 1988; Vandierendonck & de Vooght, 1993). The central executive may also provide additional processing resources to the slave systems on very memory intensive tasks (e.g. recalling long lists of digits). The central executive is often equated with general working memory factors proposed by other theorists (Daneman & Carpenter, 1980; Just & Carpenter, 1992). The Just and Carpenter theory is particularly interesting because it proposes that working memory capacity limitations arise from a limit to the total activation of units in long-term memory. Elements in working memory are thus highly-activated long-term memory units rather than elements in a separate store (Just & Carpenter, 1992). The important consideration is that elements in working memory can be considered *functionally* different from those of long-term memory rather than that they are elements of a separate store. Provided different processes operate on elements in working memory from those which operate on elements in long-term memory it makes sense to differentiate between the two systems.

Recently, Baddeley has argued that the phonological loop is comprised of two sub-components - a *phonological store* and an *articulatory control process* (Baddeley, 1986). The phonological store is a passive system which represents material in a phonological code subject to decay over time. The articulatory control process is used to maintain material in the store through rehearsal or to recode nonphonological material (e.g. written words or pictures) for entry into the phonological store. Evidence for divisions within the visuo-spatial sketchpad is less clear-cut. There is strong neurophysiological and neuropsychological evidence for a distinction between visual and spatial components of the sketchpad (Farah, 1988; Farah et al., 1988; Hanley, Young, & Pearson, 1991). Empirical work on this distinction is less convincing, though there is evidence that many tasks are more influenced by spatial than visual disruption (Baddeley & Wilson, 1988; Gathercole & Baddeley, 1993). It is not known whether visual and spatial components reflect separate sub-components of visuo-spatial working memory or differing levels of central executive involvement in visual and spatial tasks (Gathercole & Baddeley, 1993; Logie, 1990; Logie & Marchetti, 1991). One proposal is that the visuo-spatial sketchpad consists of a passive visual store and an active rehearsal process involved in visual scanning and motor control (Logie & Marchetti, 1991). Interference from 'pure' spatial tasks may therefore reflect either a separate spatial store, or some form of rehearsal process, or central executive involvement. Evidence for different components and sub-components in working memory has relied heavily on so-called interference tasks. The rationale behind these experiments is simple. If a task involves a particular resource performance on that task should be disrupted by performing a concurrent task which competes for those resources. In many cases devising 'pure' interference tasks for a particular sub-component is not trivial; as has been noted it has proven particularly challenging for sub-components of the visuo-spatial sketchpad (Baddeley, 1986; Logie, 1986; Logie, 1989; Logie, 1990; Logie & Marchetti, 1991).



## **2.2.2 Mental models and specific working memory systems**

A number of researchers have investigated the involvement of specific working memory components in the processing of mental models. Working memory theorists have argued that the central executive plays an important role in reasoning and comprehension (Baddeley, 1986; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993). In fact, some researchers have argued that reasoning ability is little more than working memory capacity (Kyllonen & Christal, 1990). Baddeley has also proposed that working memory provides a system for representing and reflecting on stored information (especially specific episodes). This, in turn, allows an organism to construct and manipulate mental models for comprehension, problem solving and prediction (Baddeley, 1992). One of the few specific proposals of the representation of mental models in working memory has been made by Glenberg and Langston. Glenberg and Langston have proposed that mental models are constructed in the spatial medium of the visuo-spatial sketchpad (Glenberg & Langston, 1992).

### **The phonological loop**

It is often assumed that the phonological loop plays little or no role in the use of a mental model for reasoning or comprehension. Baddeley and Hitch (1974) showed that articulatory suppression had a small but significant effect on verbal reasoning performance, but Toms, Morris and Ward (1993) found that neither articulatory suppression nor visuo-spatial suppression influenced performance in a study of conditional reasoning. In a study by de Vooght and Vandierendonck (1993) articulatory suppression impaired performance on one-dimensional spatial reasoning problems, while Oakhill and Johnson-Laird (1984) demonstrated that articulatory suppression slowed reading times for spatial descriptions. Morra (1989) obtained a low but significant correlation between forward digit span and recall of spatial descriptions among adults using a mental models strategy. Thus most of the available experimental evidence supports the view that the phonological loop is involved in the construction and retention of a mental model. However, the weak influence of articulatory suppression and the low correlation between digit span and retention suggest that its role is peripheral. Neuropsychological evidence reported by Baddeley and Wilson (1988) supports this view. They describe a patient with a digit span of two and a sentence span of three words who showed comprehension problems with long sentences. Baddeley and Wilson argued that the phonological store plays an important role in buffering information while a mental model of the meaning of a sentence is being set up.

### **The visuo-spatial sketchpad**

Despite the proposals of Glenberg and Langston (1992) it is not clear that the visuo-spatial sketchpad is always involved in the processing of mental models. The study of conditional



reasoning by Toms, Morris and Ward (1993) failed to find any disruption of reasoning by tapping and tracking tasks. On the other hand, studies of understanding and reasoning with spatial descriptions have shown impairment due to concurrent spatial tasks (de Vooght & Vandierendonck, 1993; Oakhill & Johnson-Laird, 1984). Unfortunately, as has already been discussed, there is also evidence to suggest that the central executive is involved in the encoding component of many supposedly spatial interference tasks (Baddeley, 1986; Gathercole & Baddeley, 1993; Logie & Marchetti, 1991; Quinn, 1991). While it is therefore plausible that spatial reasoning and comprehension utilizes the spatial component or sub-components of working memory (e.g. a passive spatial store or an active rehearsal process) the available empirical support is mixed.

### **The central executive**

By contrast, the evidence for central executive involvement in the processing of mental models is very strong. Concurrent memory load disrupts verbal reasoning performance more than articulatory suppression (Baddeley, 1986; Baddeley & Hitch, 1974). General working memory capacity is a good predictor of reasoning ability (Kyllonen & Christal, 1990) and central executive load has been shown to disrupt both conditional and spatial reasoning (de Vooght & Vandierendonck, 1993; Toms et al., 1993; Vandierendonck & de Vooght, 1993). Two experimental studies are particularly illuminating because they show strong evidence of general capacity limitations and of the use of mental models. The first, by Morra, Pascual-Leone, Johnson and Baillargeon (1991) examines the predicted working memory demands for mental model and verbal memory strategies for understanding spatial descriptions by young children (aged 9.5, 12 and 14). The working memory demands predicted by the mental models strategy fitted the observed pattern better than that of the verbatim strategy for all three age groups studied. The second, a study by de Vooght and Vandierendonck (1993) showed that performance on a one-dimensional spatial and temporal reasoning task fitted the predictions of the mental models theory better than that of a rule theory. They also demonstrated that interference tasks involving the central executive, phonological loop and visuo-spatial sketchpad all disrupted performance.

A review of the evidence suggest probable roles for all three components of the specific working memory system discussed here in the construction or utilization of a mental model. The precise role of each component is more difficult to ascertain. The phonological loop appears to play a role in the construction of a mental model from verbal or written descriptions. This role appears to be particularly important where sentences are long (Baddeley & Wilson, 1988) or otherwise difficult to comprehend, for example because of discontinuities (Oakhill & Johnson-Laird, 1984). It may also influence retention of a mental model (Morra, 1989), perhaps by freeing more time or central executive resources for encoding. The visuo-spatial sketchpad appears to be involved in spatial comprehension and reasoning, but it cannot be ruled out that this is an



artefact of central executive involvement in spatial encoding (de Vooght & Vandierendonck, 1993; Gathercole & Baddeley, 1993; Logie & Marchetti, 1991; Oakhill & Johnson-Laird, 1984). Finally, the central executive is almost certainly involved in reasoning and comprehension (Gathercole & Baddeley, 1993; Just & Carpenter, 1992; Kyllonen & Christal, 1990). It is hardly surprising that the studies reported here have found central executive involvement in a number of reasoning and comprehension tasks.

In conclusion, it appears that the central executive rather than the visuo-spatial sketchpad is likely to be the main working memory component in the processing of mental models. The phonological loop may be indirectly involved in the process of mental model construction. There is as yet little experimental evidence to link the visuo-spatial sketchpad with the construction and utilization of a mental model, although there is evidence that a spatial sub-component of the sketchpad is involved in reasoning and comprehension with spatial descriptions.

## 2.3 Experimental studies of mental models in working memory

Early experimental evidence for mental models tended to rely heavily on studies of long-term memory (Bransford et al., 1972; Ehrlich & Johnson-Laird, 1982; Garnham, 1981; Garnham et al., 1982; Mani & Johnson-Laird, 1982). Even research on reasoning has based its findings more often on the product of a reasoning task than the process (Johnson-Laird, 1983; Johnson-Laird & Steedman, 1978). There are good reasons why the representation of mental models in working memory has been neglected. Firstly, studying the processes of comprehension and reasoning 'on-line' is difficult.<sup>4</sup> Secondly, the characteristics of long-term memory for mental models can be informative about representation of mental models in working memory. Unfortunately, though studies of long-term memory can provide useful information about the representation of mental models in working memory they can also be misleading.

Recently, there has been a shift towards greater interest in the way mental models are manipulated and updated in working memory. A number of psychologists and psycholinguists have begun to carry out experiments to investigate mental models in the processing of discourse on-line (Bower & Morrow, 1990; Glenberg et al., 1987). Several experiments have used priming of spatial location as a methodology for studying narrative comprehension (Bower & Morrow, 1990; McNamara, 1991; McNamara, 1992). Other studies of the on-line comprehension of discourse

---

<sup>4</sup> For the purposes of this thesis 'on-line' processes are those that rely almost exclusively on working memory, though the possibility that 'informationally encapsulated' lower-level systems may be involved is not discounted (Fodor, 1983).



have implicated mental models in anaphoric reference (Garnham, 1987; Garnham & Oakhill, 1992; Speelman & Kirsner, 1990) and inference generation (Ehrlich & Tardieu, 1992; Garnham & Oakhill, 1992; Noordman & Vonk, 1992; Tardieu, Ehrlich, & Gyselinck, 1992).

Research into spatial mental models in on-line narrative comprehension shares a number of assumptions. Firstly, it assumes that elements in working memory, such as tokens in a mental model, may possess different levels of activation (Glenberg & Langston, 1992; Glenberg et al., 1987; Morrow, Bower, & Greenspan, 1989; Morrow, Greenspan, & Bower, 1987). Secondly, it assumes that tokens in a mental model can 'point' or refer to other elements in working or long-term memory (Glenberg & Langston, 1992). Finally, it assumes some mechanism for 'foregrounding' or 'focusing' attention on key elements in the narrative (Bower & Morrow, 1990; Glenberg et al., 1987). Bower and Morrow (1990) illustrates the notion of focus or foregrounding with the 'spotlight' metaphor. In narrative comprehension the shifting spotlight follows the protagonist's path through space and time (Bower & Morrow, 1990; O'Brien & Albrecht, 1992).<sup>5</sup> Objects, characters and events in the foreground (the 'here and now' of the narrative) are at higher levels of activation, can be more readily accessed and are more likely to prime (e.g. spatially) associated items. Glenberg found evidence that mental models contribute to the foregrounding process in text comprehension (Glenberg et al., 1987). Items that are in spatial proximity to the protagonist (who is assumed to be foregrounded) are more quickly accessed than items not spatially associated with the protagonist.

Although many of the findings have been primarily concerned with the representation of goals, plans, motives or emotions in mental models of narratives (Bower, 1989; Garnham & Oakhill, 1992; Gernsbacher, Goldsmith, & Robertson, 1992) some of the most interesting results have related to the spatial representation of narratives as the study by Glenberg et al. (1987) illustrates. Morrow, Greenspan and Bower (1987) gave people maps of buildings to learn and then asked them to read stories set in those buildings. Each map location contained one or more objects unique to that particular room, and each story had a new character with a distinct goal. They found that subjects were quicker to respond to object probes from goal rooms (where they were heading) than source rooms (the prior location of the character). Responses were slower still for other rooms in the building and slowest of all to rooms outside the building. This was true even when the source room was most recently mentioned. Incidental mention of locations in passing did not change focus provided the location was irrelevant to the current actions of the protagonist. Morrow, Bower and Greenspan (1989; 1990) investigated whether intermediate locations or landmarks experience activation or whether only explicitly mentioned locations are activated.

---

<sup>5</sup> Presumably foregrounding follows different (but related) principles for other forms of text. For example, perhaps expository texts which offer thematic or conceptual protagonists make foregrounding and hence comprehension easier.



Intermediate locations are those that a character has to pass through to travel from the source room to the goal room, but which are not explicitly mentioned in the text. They found that responses to objects from the goal room were fastest, followed by those from the intermediate location, with those from source rooms slower still. These findings have since been replicated, with minor reservations. Readers only construct highly detailed spatial mental models when task demands (i.e. the kind of object probe used) require them to follow the protagonist through the learned spatial layout (Wilson, Rinck, McNamara, Bower, & Morrow, 1993). It is also worth noting that these highly detailed spatial mental models are constructed with the aid of spatial information retrieved from long-term memory.

Where people adopt a point of view where the observer is embedded within the scene, or where people conceptualize more than one two-dimensional scene different kinds of spatial mental models are probably constructed (Franklin, Tversky, & Coon, 1992; Taylor & Tversky, 1992a; Taylor & Tversky, 1992b; Tversky, 1991). An important consideration that has emerged from these and other studies is that people can and do conceptualize space in different ways; responding to the spatial demands of the situation and the task (Franklin et al., 1992; O'Brien & Albrecht, 1992; Tversky, 1991; Wilson et al., 1993). Understanding how these demands influence the representation of spatial mental models in working memory may illuminate the interpretation of evidence for their representation in long-term memory. In many cases the spatial mental models people have investigated may also involve switching information between working memory and long-term memory. For both these reasons the question of the relationship between working memory mental models and what is represented in long-term memory is crucial.

## **2.4 Mental models and long-term memory**

It has already been suggested that the representation of mental models in working memory relies heavily on studies of long-term memory. In the following sections this research is reviewed and conclusions are drawn about the status of mental models and in particular spatial mental models in long-term memory.

### **2.4.1 Memory and reasoning with mental models**

In studies of reasoning with mental models the nature of this representation is often only implicated in so far as it motivates the predicted pattern of reasoning performance (Byrne, Handley, & Johnson-Laird, 1992; Byrne & Johnson-Laird, 1989; Johnson-Laird & Byrne, 1989; Johnson-Laird & Byrne, 1990; Johnson-Laird, Byrne, & Tabossi, 1989; Oakhill & Johnson-Laird, 1985b; Oakhill et al., 1989). However, the evidence that people reason using mental models in



many situations is quite strong (Baron, 1988; Evans et al., 1993; Eysenck & Keane, 1990; Gilhooly, 1988; Johnson-Laird & Byrne, 1991), and relies on people remembering and reporting their conclusions. Experiments by Byrne and Johnson-Laird also suggest that people can remember the intermediate conclusions predicted by the mental models account (Byrne & Johnson-Laird, 1990; Johnson-Laird & Byrne, 1991). They gave people a surprise recognition test for the conclusions they had drawn in a reasoning task, but eliminated the correct 'no valid conclusion' response from the set of recognition items. Of the people who had originally reached the correct conclusion 20 percent realized it wasn't on the list, 6 percent were unable to decide and the remaining 74 percent all chose the conclusion that matched the predicted initial mental model of the premises. This is very difficult to account for by competing rule-based theories of reasoning (because there are no logical proofs for invalid problems). It shows that people may be able to remember not only information about the final model constructed but also information from intermediate steps or models in the reasoning process.

## **2.4.2 Mental models and memory for text**

The classic studies of gist memory by Bransford and colleagues are frequently cited as support for the view that people represent and remember the situation described by discourse and that this representation takes the form of a mental model (Bransford et al., 1972; Bransford & Franks, 1971; Glenberg et al., 1987; Johnson-Laird, 1989). Using a similar paradigm, Garnham has shown that people tend to recall mental models of events rather than linguistic expressions describing those events (Garnham, 1981).

Further evidence comes from a study of the effects of different reading goals on memory for text. Schmalhofer and Glavanov demonstrated that readers asked to summarize a manual remembered more propositional information while readers asked to learn the whole text remembered more situational information (Schmalhofer & Glavanov, 1986). They also noted that situational information was accessed faster by all readers and (as in work by Bransford) influenced sentence recognition more strongly than propositional information. Reading for text summarization took longer and was less likely to facilitate the construction of an elaborate mental model. Reading for learning may have been quicker because it relied more on the reader's domain knowledge. This explanation would match the findings of research on expertise which suggests that experts construct richer, more elaborate mental models which facilitate inference generation during narrative comprehension (Noordman & Vonk, 1992; Tardieu et al., 1992).

A study of recognition memory by Fletcher and Chrysler also provides convincing evidence of the distinction between mental models and representations of the language of a text (Fletcher & Chrysler, 1990). They tested recognition memory for surface form, propositions and



situations derived from the text by varying similarity between the targets and distracters. Recognition performance was best when distracter sentences were dissimilar in terms of their surface form, propositional content and situational content. Recognition got worse as the similarity of the distracters to the original sentences increased. Because information about surface form, propositional content and the situation described by the text all influenced recognition they argued that this supported the distinction between mental models and representations of the language of a text.

Using a different task, Glenberg and Langston have argued that pictures and diagrams facilitate the construction of mental models. Pictures can only do this, however, if they depict the structural relationships of the situation described by the text (Glenberg & Langston, 1992). They presented subjects with a series of texts describing a procedure made up of several stages. All of the stages, except the middle two, have to be carried out in a strict order. The middle two stages could be carried out in either order. The structure of the text and the structure of the procedure thus differed. The pictures only facilitated memory for the procedures if they correctly depicted the structure of the procedure. In a control condition where text order did not differ from the order of the procedure the pictures failed to improve recognition memory. This cannot be explained by the influence of repetition, motivation or simple versions of dual-coding theory because pictures do not always aid retention (Glenberg & Langston, 1992; Paivio, 1986).

Studies of reasoning and memory for text both suggest that mental models are remembered. However, there is little or no evidence from this research on the characteristics of this long-term memory representation. It is very difficult to identify the underlying representation where more than one type of representation may be involved in a given task. Conclusions derived from a mental model might be translated into a verbatim representation for verbal report (though this would not explain the findings presented in section 2.4.1). Similarly, gist recall might be supported by a propositional representation of the situation rather than a propositional representation derived from the language of the discourse (Kintsch, 1988). The remaining discussion looks in more detail at research on memory for mental models, and especially spatial mental models, in order to clarify some of these issues.

### 2.4.3 Mental models and the fan effect

The *fan effect* discovered by Anderson (1984; 1973) has long been used to support the view that memory is structured in the form of a propositional network. The fan effect is “an increase in retrieval time or error rate ... accompanying an increase in the number of newly learned associations for a concept” (Radvansky & Zacks, 1991). Radvansky and Zacks have proposed a different explanation of the fan effect based on the notion of mental models (Radvansky,



Spieler, & Zacks, 1993; Radvansky & Zacks, 1991). They argue that where new associations for a concept can be integrated into a single, coherent mental model (such as when all the associated concepts are consistent with a single spatial location) no fan effect will arise. They demonstrated that if sentences like those in Figure 2.2 were learned a fan effect will occur for sentences 1, 2 and 3 but not 4, 5 and 6 (compared with retrieval time or errors for sentence 7):

<u>Multiple location sentences:</u>	<u>Single location sentences:</u>
1. The cola machine is in the hotel.	4. The display case is in the city hall.
2. The cola machine is in the public library.	5. The potted palm is in the city hall.
3. The cola machine is in the high school.	6. The broken window is in the city hall.
<u>Control sentence:</u>	
7. The welcome mat is in the barber shop.	

**Figure 2.2: Examples of the materials used by Radvansky and Zacks (1991).**

It is argued that people organize long-term memory around structures of real world situations and that the fan effect is caused by interference from irrelevant mental models. Radvansky and Zacks also found no evidence of a fan effect within a single mental model but speculate that there may be an effect for larger fan sizes (Radvansky & Zacks, 1991).

Organization of the mental models constructed in these experiments does not appear to be influenced by the transportability of the objects or by the use of the indefinite rather than definite article (Radvansky et al., 1993). Including animate entities (such as a person) in the sentences influenced some, but by no means all, people to adopt person-based rather than location-based mental models consistent with evidence from other studies of spatial mental models (Bower & Morrow, 1990; Glenberg et al., 1987; O'Brien & Albrecht, 1992; Radvansky et al., 1993; Tversky, 1991). These experiments by Radvansky and her associates go against traditional propositional accounts of organization in long-term memory. While propositional theories could account for these results, to do so they would need to acknowledge (at least implicitly) the legitimacy of mental models (Radvansky & Zacks, 1991). This is because the most important organizational element of the theory would be a mental model (e.g. made up of a set of propositions) rather than a single proposition.

**2.4.4 Referential continuity and the construction of spatial mental models**

Referential continuity is known to be an important factor in text comprehension and several researchers have proposed that this is best accounted by its role in the construction of a mental

model (Garnham et al., 1982; Johnson-Laird, 1980; Johnson-Laird, 1981). Ehrlich and Johnson-Laird (1982) provide an excellent demonstration of the importance of referential continuity, but also show that the importance of referential continuity does not rely on a linguistic representation but on a semantic representation of discourse. In their first experiment they showed that discontinuous spatial descriptions (where referential continuity is not maintained between sentences) are harder to recall than continuous spatial descriptions. Their remaining two experiments used three types of spatial description discontinuous, semicontinuous and continuous (see Figure 2.3).

<u>Continuous description:</u>	<u>Discontinuous description:</u>
The calf is behind the deer	The hammer is on the left of the pins
The deer is on the left of the goat	The ruler is on the left of the paper
The rabbit is in front of the goat	The paper is in front of the hammer
<u>Semicontinuous description:</u>	
The bread is in front of the chocolate	
The chocolate is on the left of the meat	
The cheese is on the left of the bread	

**Figure 2.3: Examples of materials used by Ehrlich and Johnson-Laird (1982).**

Semicontinuous descriptions are those in which referential continuity between sentences is not maintained but where new sentences always refer to items which should already be present in a spatial mental model. Thus the mental model account predicts that semicontinuous descriptions should be easier than discontinuous descriptions. A propositional or linguistic account of spatial representation predicts that semicontinuous descriptions should be as difficult to understand and remember as discontinuous ones. The results showed that readers did find discontinuous descriptions harder to remember than semicontinuous descriptions, and that there was no significant difference in retention between continuous and semicontinuous descriptions.

Ehrlich and Johnson-Laird (1982) argued that difficulty with the discontinuous descriptions arise from having to hold two independent relations (each containing two items) in working memory while the third sentence is read or because the discontinuous descriptions are represented in a linguistic or propositional code that is harder to recall. A subsequent study replicated these findings and concluded that discontinuous descriptions suffered greater disruption from articulatory suppression suggesting that some readers switch to encoding



relations in a linguistic or propositional form (Oakhill & Johnson-Laird, 1984). Experiments by Morra have also replicated these findings with adults and older children and underlined the role of working memory in the construction of spatial mental models (Morra, 1989; Morra et al., 1991). It seems that people construct mental models to understand spatial descriptions and that referential continuity between new relations and items already in the mental model (rather than referential continuity between sentences) aids in the mental model construction process. Note that sentences which are more easily integrated into a spatial mental model are more easily remembered. As will be discussed later in this chapter, Johnson-Laird appeals to notions of greater amount of processing, greater depth of processing and greater elaboration to account for the difference in memorability between mental models and propositional or linguistic representations ( Craik & Lockhart, 1972; Craik & Tulving, 1975; Johnson-Laird, 1983; Mani & Johnson-Laird, 1982).

### 2.4.5 The determinacy of spatial descriptions; propositions versus mental models

A classic experiment by Mani and Johnson-Laird (1982) manipulated the determinacy of spatial descriptions in order to investigate how people represent and remember spatial information. Propositional representations are able to explicitly represent both indeterminate and determinate descriptions (Palmer, 1978; Rumelhart & Norman, 1983). Mental models, on the other hand, have difficulty in representing indeterminacy explicitly (see sections 2.2.1 and 2.1.4) Mani and Johnson-Laird gave subjects determinate or indeterminate spatial descriptions to read (see Figure 2.4). The linguistic content of the two types of description is very similar. In fact, they differ only by the substitution of a single relation term. However, changing a single spatial relation is capable of rendering a description *radically indeterminate*. Thus, even though the indeterminate description consists of four determinate spatial relations the description as a whole can support the construction of at least two distinct spatial mental models. This is the sense in which they are radically indeterminate (it is not merely a question, for instance, of uncertainty about the distance between items).

<u>Determinate description:</u>	<u>Indeterminate description:</u>
The hospital is <i>behind</i> the garage	The hospital is <i>in front of</i> the garage
The prison is in front of the garage	The prison is in front of the garage
The garage is to the right of the hotel	The garage is to the right of the hotel
The theatre is to the left of the hospital	The theatre is to the left of the hospital

**Figure 2.4: Example materials from Mani and Johnson-Laird (1982).**



The initial task of the participants was to decide whether diagrams they were given matched the original description or not. Later they were given a surprise recognition test requiring them to rank four items (the original description, an inferable one describing the same configuration of objects and two foils). A similar recognition test to that used by Mani and Johnson-Laird is described in more detail in the introduction to Chapter 4. Mani and Johnson-Laird found that verbatim memory (exhibited when the original was ranked higher than the inferable) was higher for indeterminate descriptions. Gist memory (exhibited when the inferable and the original were both ranked higher than the foils) was higher for the determinate descriptions. This finding was interpreted as evidence for the construction and retention of mental models. Determinate descriptions permit the construction of a single mental model which supports gist recall. For indeterminate descriptions construction of a single consistent mental model is not possible, therefore people abandon their attempt to construct a mental model and concentrate on remembering the description itself. Johnson-Laird (1983) has argued that this crossover interaction cannot be explained without postulating at least two sorts of mental representation.

Payne (1993) has questioned the interpretation of Mani and Johnson-Laird's findings. In order to consider Payne's alternative it is necessary to reconsider the purpose of the original experiment. Mental models are analogical representations formed in working memory for reasoning and comprehension. However, the Mani and Johnson-Laird experiment attempts to provide evidence for mental models by investigating how spatial descriptions are represented in long-term memory. The need for a distinction between the long-term memory representation of mental models and the representation of mental models in working memory has already been discussed. Payne argues, contrary to Mani and Johnson-Laird, that mental models are not represented in long-term memory (or at least not analogically). Instead he proposed that what people retain in long-term memory are the operations used to construct the mental model, rather than the model itself. These operations are recorded as a set of propositions in long-term memory which Payne calls an *episodic construction trace*.

Payne attempted to replicate Mani and Johnson-Laird with more strictly controlled materials. The foils used by Mani and Johnson-Laird in the recognition test described spatial layouts of a different shape to those in the original descriptions. Payne constructed foils describing the same shape as the original and inferable descriptions, eliminating the possibility that people were only remembering the shape rather than a spatial mental model. Payne also ensured that all the descriptions in the surprise recognition test shared only a single common sentence. Finally, a stricter scoring procedure for recognition memory was used. A response was only scored as verbatim memory if the original description was ranked first. However, he was unable to find evidence that subjects were remembering mental models. Instead, his results suggested that memory was dependent on the *overlap* between the process of mental model



construction for the original description compared with the descriptions presented in the surprise recognition test. This overlap (henceforth termed *trace overlap*) is operationalized as the number of shared propositions in the *episodic construction trace* of the two descriptions.

Payne (1993) reimplemented the computer model described by Johnson-Laird (1983) in order to make the predictions of his theory more explicit. In both cases the computer program constructs arrays in working memory. Unlike the version written by Johnson-Laird, however, the version by Payne discards the array it has constructed in working memory. It stores only a record of the operations involved in constructing the spatial mental model. The episodic construction trace records operations in the form of a list. The first element in the list is usually a token already present in the array, the second token is usually a token being added to the array, and the third token usually describes how the new token was located relative to the first token. So the list [table chair front] records the operation of placing the token 'table' in front of the token 'chair' (which is already present in the array). There are two exceptions to this general procedure for recording the operations used to construct the array:

1. The insertion of the initial pair of tokens in the array is denoted by the symbol 'start' as a modifier to the operation. This indicates that two new tokens have been added to the array.
2. If an operation can not proceed because it would render the description indeterminate the symbol 'clash' is recorded as a modifier to the operation that produced the clash. In addition a sub-list containing the tokens involved in the 'clash' is recorded as the final element of the list.

The episodic construction trace is stored propositionally in long-term memory. In the computer model this takes the form of an unordered list. The list is unordered in the sense that each separate proposition is equally accessible in long-term memory. It is important to note that the episodic construction trace preserves information about the order of spatial mental model construction. However, some (but not all) information about the order of construction is implicit in the operations which have been recorded, rather than being explicitly preserved in the order or accessibility of the propositions in long-term memory. Partial forgetting is simulated by a fixed percentage chance that a random list will be deleted. Figure 2.5 shows the episodic construction trace for the two descriptions in Figure 2.4. Note that despite the change of only one word there are only two shared propositions (the trace overlap) between the descriptions.

Payne found that trace overlap predicted recognition performance better than verbatim or mental model accounts. He argues that the findings of Mani and Johnson-Laird (1982) were due to an experimental artefact. It is possible to construct inferable determinate descriptions with a trace overlap of 3 (i.e. inferable descriptions which share three propositions in their respective episodic construction traces). For indeterminate descriptions the maximum trace overlap for a corresponding inferable description is 1. Thus the Mani and Johnson-Laird crossover interaction

could be a product of this ‘psychological distance’ between the inferable and original description rather than genuine memory for mental models.

<u>Determinate episodic construction trace:</u>	<u>Indeterminate episodic construction trace:</u>
[Start hospital garage behind]	[Start hospital garage front]
[ prison garage front]	[ prison garage front [clash prison hospital]]
[hotel garage left]	[hotel garage left]
[theatre hospital left]	[theatre hospital left]

**Figure 2.5: Episodic construction traces for the descriptions in Figure 2.4.**

In a stronger test of the episodic construction trace model Payne (1993) showed that reordering the sentences in the original descriptions depressed recognition scores. Reordering disrupts the episodic construction trace but should have no effect on either the individual sentences or the array that is described. For this reason it is difficult to argue that recognition depends on either verbatim memory for the description or memory for the mental model itself. One possible problem for the episodic construction trace account is that it seems counter-intuitive that people do not remember mental models they construct, but only the operations used to construct the models. A number of the studies of spatial memory and reasoning already discussed suggest that people can and do remember the mental models they construct. The next two sections look at studies of memory for spatial descriptions and narratives which apparently contradict the finding that people only remember the process of spatial mental model construction and not the spatial mental model itself.

**2.4.6 Perspective, order and memory for spatial mental models**

In a series of experiments Holly Taylor and Barbara Tversky asked subjects to learn stories incorporating spatial descriptions of places (Taylor & Tversky, 1992a; Tversky, 1991). Later, subjects were given a true/false sentence verification task. They were also asked to draw the environment described in each story. Half the stories took a ‘route’ perspective and half a ‘survey’ perspective. A route perspective is characterized by a ground-level, moving viewpoint; directions are usually relative to the protagonist of the narrative. A survey perspective is characterized by a stationary viewpoint from above; directions are usually canonical in form, as if describing a map (Thorndyke & Hayes-Roth, 1982). Other researchers have shown that the perspective taken by a spatial narrative influences the mental model that readers construct (Perrig & Kintsch, 1985). However, text perspective may have been confounded with how well the texts were organized (Taylor & Tversky, 1992a). Tversky and Taylor equated the coherence



of the narratives they used by asking pilot subjects to rate them and they were also able to show that both 'route' and 'survey' texts were equally well remembered.

They found that verbatim statements (sentences taken from the original text) were verified more quickly than inference statements (statements that were true with respect to the situation described in the story but which were not explicitly mentioned in the text). This provided evidence that readers of the stories had some form of mental representation of the language of the text. The inference statements in the verification task adopted either a route or a survey perspective. Tversky and Taylor found that the perspective of the inference statements did not influence verification accuracy or speed. They argued that this required a 'perspective-free' representation of the situation described by the text. Hence, this mental representation was like a spatial mental model rather than image (Johnson-Laird, 1983). The nature of this spatial mental model is reminiscent of the spatial representations of people navigating very familiar environments (e.g. refer to the study by Thorndyke and Hayes-Roth described section 1.2.2).

When Tversky and Taylor asked people to draw the environments they had read about it was noted that the drawing order of the features in the description was significantly related to the order items were mentioned in the description (Taylor & Tversky, 1992a; Tversky, 1991). According to Tversky (1991) this drawing order effect arises because participants in their studies "reconstruct the mental model of the environment in the same order as they originally constructed it, i.e. in the order of the description they read". Furthermore she argued that:

"... readers' mental models are not image- or map-like. If they were, there should be no differences in drawing order depending on description perspective; rather, drawing order should depend on the characteristics of the image or map alone ... "

[Tversky (1991) p. 130]

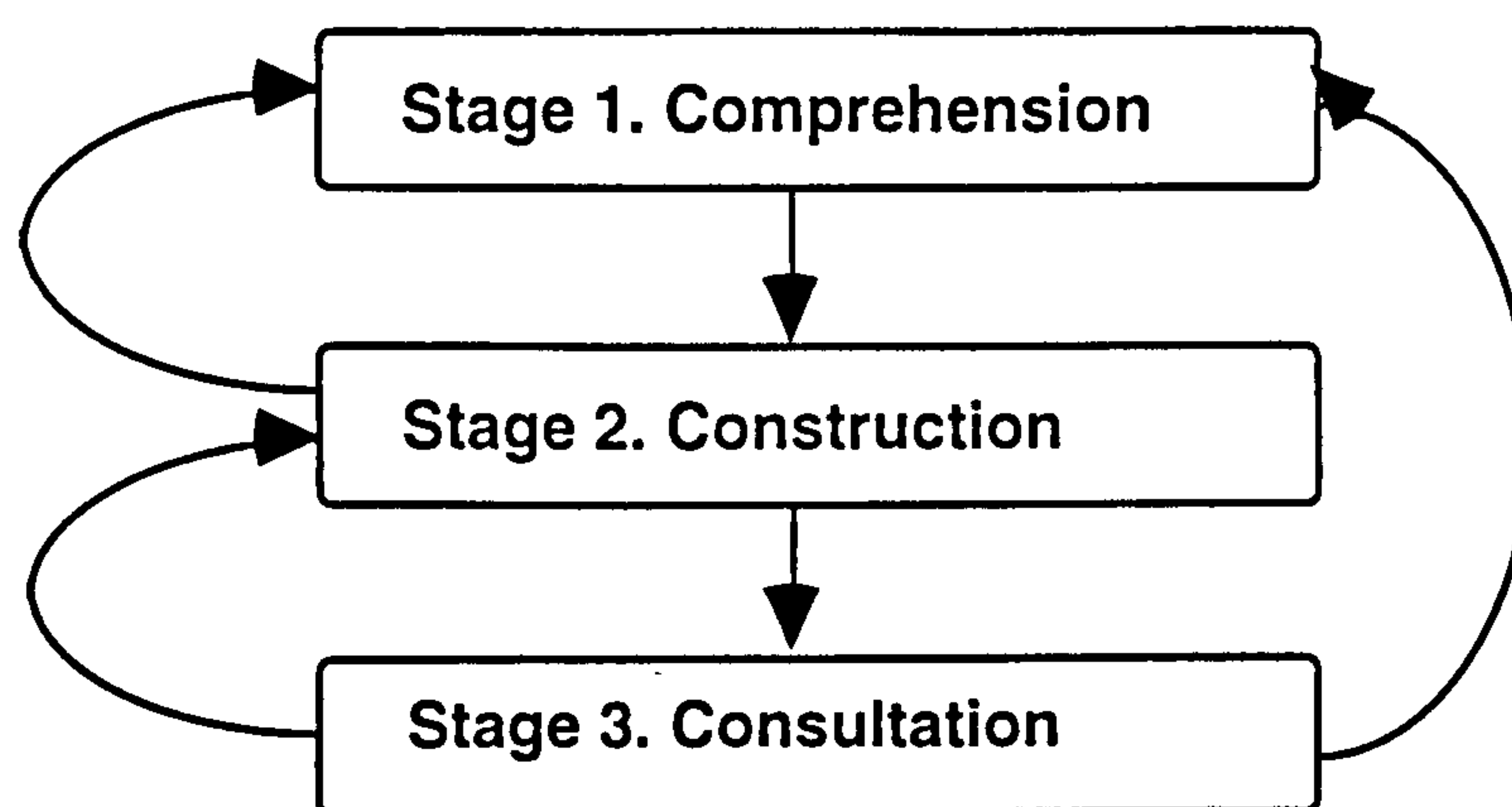
In other words an image would not preserve structural information about the order of construction. On the other hand a spatial mental model is a set of structural relationships between entities, therefore it can, in principle, encode structural information about the order or process of mental model construction (Johnson-Laird, 1983; Marr, 1982; Palmer, 1975; Tversky, 1991). It should be noted though that Tversky herself does not provide any indication why such information should necessarily be preserved. It appears then, that Tversky and Taylor have provided evidence both that the structure of spatial mental model can be remembered and that ordinal information about the process of mental model construction may be preserved in the representation of a spatial mental model in long-term memory. For a more critical discussion of these conclusions the reader is referred to the introduction to Chapter 3.

## 2.5 What constitutes memory for a spatial mental model?

A number of issues have emerged from the previous discussion. Studies of long-term memory are often used to provide evidence for the thesis that people construct, update and manipulate mental models in long-term memory. Unfortunately this form of argument raises the possibility that long-term memory structures will be confused with structures in working memory. The approach adopted here, is to provide a theoretical account of the relationship between mental models in working memory and the resulting long-term memory representation. Such an account needs to be broadly consistent with existing mental model theories, and, most importantly of all, must provide predictions for empirical research. The next section outlines a 'working model' of the construction of spatial mental models for spatial descriptions.

### 2.5.1 An outline of spatial mental modeling

The process described below is based on a computer program for interpreting spatial descriptions written by Johnson-Laird (1983). Johnson-Laird has used the program to demonstrate that his theory does not capitalize on a 'visual' metaphor. The procedures embodied in the program are able to construct and generate conclusions from a mental model without being able to 'see' the arrays that make up the model (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). It is this program which Payne has modified to implement his *episodic construction trace* account of memory for spatial descriptions (Payne, 1993).



**Figure 2.6: The three stages in the cycle of spatial mental modeling.**

Other details of the process are drawn from by Johnson-Laird and Byrne (1991). For the sake of simplicity, mental models, including spatial mental models, are assumed to be constructed within a general working memory system such as the central executive (see section 2.2). The processes described here are limited to the construction of mental models from spatial descriptions, but would form the basis for the construction of spatial mental models from



perception, from long-term memory or, as in many everyday situations, from more than one source. The outline has been divided into three separate stages for clarity. The three stages are described in the order which they first occur. However, it should be noted that spatial mental modeling is a cyclical, ongoing activity (see Figure 2.6). Once the first sentence of a description has been read, appropriate tokens and relations entered into the model, the process begins again with the next sentence. In most, if not all cases, construction is followed by some form of consultation process to check the model is consistent with the original description.

### **Stage 1. Comprehension**

Before construction of a spatial mental model can begin it is necessary to read and understand the meaning of the spatial relationships contained in the description. The meaning of each spatial relationship is derived from the lexical meaning of linguistic relations but is also influenced by general knowledge. Mental model construction occurs on-line and so demands on verbal or phonological working memory appear to be an important factor in the construction process, especially when people attempt to construct mental models from difficult sentences, for example those involving a discontinuity (Baddeley & Wilson, 1988; Ehrlich & Johnson-Laird, 1982; Morra, 1989; Oakhill & Johnson-Laird, 1984). Sentences describing difficult relations take longer to read (Ehrlich & Johnson-Laird, 1982; Oakhill & Johnson-Laird, 1984) and it is suggested that they are rehearsed while an attempt is made to integrate the new spatial relationships into the mental model. With complex spatial descriptions the mental model construction process is not only slowed down but may even be abandoned in favour of a representation closer to the linguistic form of the description. Comprehension processes are also involved in checking or refreshing information in the model.

### **Stage 2. Construction**

The spatial mental model described here consists only of tokens and spatial relationships between tokens. However, other relationships such as temporal or causal relationships are possible and would follow the same general procedure outlined here (de Vooght & Vandierendonck, 1993; Schaeken et al., 1993). Tokens and relations are assumed to possess an activation level. Activation of tokens and relations decay with time, but may increase if a given relation or token in the model is refreshed or inspected. Tokens and relations in the model also increase in activation when adjacent tokens are added, inspected or refreshed. Tokens may refer or point to other elements in working memory or long-term memory. These would be elements strongly associated with the tokens in the spatial mental model (Bower & Morrow, 1990; Glenberg et al., 1987). Elements that are pointed or referred to increase activation in proportion to the strength of the association and the activation level of the token in the mental model.

Construction of a spatial mental model begins after a novel spatial relationship is read and comprehended:

1. Two tokens corresponding to the initial spatial relationship given in the description are placed in the mental model in two separate steps. The order in which these tokens are entered into the model is determined by the symmetry or asymmetry of the spatial relationship being considered:<sup>6</sup>
  - a. Where the initial spatial relationship in the description is asymmetric the second item is entered as the first token in the model.
  - b. Where the initial spatial relationship in the description is symmetric the items are entered into the model in the order they occur in the description.
2. The next step in the spatial mental model construction process depends on the items referred to in subsequent spatial relations:
  - a. If a subsequent spatial relation refers to a single novel item a new token is entered into the model relative to the positions of existing tokens.
  - b. If a subsequent spatial relation refers to two tokens already present in the model they are inspected and it is determined whether or not the new relation is consistent with the existing spatial mental model.
  - c. If a subsequent spatial relation refers to two items not represented in the current spatial mental model then (subject to working memory limitations) a new spatial mental model is begun as described in Step 1a or 1b.
  - d. If a subsequent spatial relation refers to two existing tokens which are not in the same spatial mental model then the two representations are integrated into a single spatial mental model.
3. If at any point in the construction process an apparent conflict arises between a new spatial relation and an existing spatial relation one or more of the following procedures is executed:
  - i) The content of the spatial mental model is checked with respect to the content of the spatial description (e.g. to determine whether a sentence was misread or misinterpreted).

---

<sup>6</sup> Asymmetric spatial relations are those like 'X is to the right of Y' or 'P is behind Q' that do not retain the same meaning if reversed. A symmetric relation retains the same meaning when reversed (e.g. 'A is opposite B' means the same as 'B is opposite A'). Most common spatial relations are asymmetric.



- ii) An attempt is made to reinterpret the new spatial relation, the existing spatial relation or both in order to eliminate the conflict. Typically this will involve revising an interpretation of a spatial relation (e.g. abandoning an assumption that two objects cannot occupy the same spatial location).
- iii) The spatial mental model is inspected using a recursive procedure (Johnson-Laird, 1983) to determine whether the tokens in the model can be updated or whether the model can be reconstructed to resolve the conflict.<sup>7</sup>
- iv) The spatial mental model construction process is abandoned.

Note that, where possible, tokens are placed into the model with reference to a token already present in the model (Johnson-Laird, 1983; Payne, 1993). The steps described above capture many of the procedures detailed by Johnson-Laird (1983) with one substantive addition. The placing of the two initial tokens in the spatial mental model has been broken down into two separate operations (see Figure 2.7 below).

<u>Sentence being read</u>	<u>Operation being performed</u>	<u>Spatial mental model</u>	
The city is in front of the mountain.	1. Enter mountain in model.	Mountain	
	2. Place city in front of mountain.	Mountain	City
The mountain to the left of the lake.	3. Place lake to right of mountain.	Mountain	Lake
		City	
The forest is in front of the lake.	4. Place forest in front of lake.	Mountain	Lake
		City	Forest

**Figure 2.7: The process of mental model construction for a simple spatial description**

---

<sup>7</sup> The program described by Johnson-Laird (1983) always checks for other possible models whether an inconsistency is present or not. However, as the process outlined here is a general one, not specifically intended for inference generation, this procedure is assumed to be carried out only if an inconsistency has been detected.

While this interpretation of the construction process is not explicit in the accounts of either Johnson-Laird or Payne it is consistent with the principle implicit in both accounts that if possible new tokens are always added to a spatial mental model in relation to existing tokens (Johnson-Laird, 1983; Payne, 1993). The importance of the symmetry of the spatial relation can be demonstrated by an illustration of the operations necessary to construct a spatial mental model with asymmetric spatial relations (see Figure 2.7).

### Stage 3. Consultation

Spatial mental models are constructed to understand spatial situations and to use that understanding to communicate information, to make predictions, to plan actions and to solve problems. Spatial mental models do this by supporting a variety of 'consultation processes'. Consultation processes are also involved in the construction of the mental model and are essential in order to be able to integrate new information into the representation. The simplest form of processing that the spatial mental model supports is the 'checking' of information in the model; this, and other consultation process, may serve to 'refresh' relations in the spatial mental model which, in turn, prevent loss of elements in the model through decay. Consultation processes often take the form of simple procedures such as counting and matching (Newell, 1990). It may be that consulting a spatial mental model is synonymous with generating a view of that model in the form of a visual image, though this has not been strongly advocated (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). When a part of a spatial mental model is consulted (for instance when an element is updated) it is possible that new relationships will be 'noticed'. Glenberg and Langston suggest that noticing new relationships in this way forms the basis of inference generation and is an important determinant of memorability in a mental model (Glenberg & Langston, 1992). Consultation processes are thus important for several reasons. Firstly, they are essential in monitoring or checking the construction of the spatial mental model. Secondly, they serve to maintain activation of tokens and relations and hence to 'refresh' the contents of the spatial mental model. Thirdly, they allow information, such as inferred spatial relations, to be accessed from the spatial mental model.

## 2.5.2 How many ways can a spatial mental model be remembered?

Earlier it was suggested that an account of the processes involved in the construction of a spatial mental model would help clarify our understanding of what constitutes memory for a spatial mental model. According the outline given above spatial mental modeling involves three relatively distinct stages of cognitive processing. The first stage of processing has been called *comprehension* and involves grasping the meaning of the description being read or heard (Johnson-Laird & Byrne, 1991). The second stage is the *construction* of the spatial mental model itself (Johnson-Laird, 1983; Payne, 1993). The third stage involves one or more consultation processes to check, match or otherwise read off information from the model (Glenberg &



Langston, 1992; Johnson-Laird & Byrne, 1991; Newell, 1990). All three stages of processing act in working memory, but are capable of influencing what is stored in long-term memory. It follows that memory performances attributed to the construction of a spatial mental model could depend on one or more of these stages of processing. What form of long-term memory trace would each stage of processing be associated with?

### Memory for comprehension processes

In the simple case being considered - that of comprehending spatial descriptions - memory for comprehension processes consists of memory for the language of the description. It is possible to distinguish between at least two different levels of representation of the language of the description; *surface form* and *propositions* (Fletcher & Chrysler, 1990; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Surface form representations preserve the exact wording of the discourse. Propositional representations preserve the meaning and overall structure of the text but not its exact wording. For simple descriptions it is likely that the difference between propositions and surface form is relatively small. Memory for the language of the text is likely to be most strongly influenced by the number of times it is read, heard or rehearsed. While comprehension processes are essential to spatial mental model construction in themselves they can provide no evidence of memory for a spatial mental model. This is because retention of the language of the text is possible even if a spatial mental model is not constructed. In fact, before memory performance can be attributed to memory for a spatial mental model the possibility that performance is based on a linguistic representation has to be discounted (Fletcher & Chrysler, 1990; Mani & Johnson-Laird, 1982; Payne, 1993; Tversky, 1991).

### Memory for construction processes

It is possible to remember the operations carried out to construct a particular spatial mental model. The *episodic construction trace* proposed by Payne (1993) takes exactly this form. Payne proposes that this takes the form of an unordered sequence of propositions, however it is also possible that other information about the construction process is retained. For instance there is evidence that the order of spatial mental model construction is preserved (Taylor & Tversky, 1992a). This ordinal information could be preserved in the content of an ordered episodic construction trace or it could be a side effect of the construction and consultation processes. For instance tokens integrated into a mental model early on are more likely to receive activation or to be 'noticed' and refreshed later in the construction process. Evidence from experimental studies suggests that these tokens would also receive activation when adjacent tokens are updated (Bower & Morrow, 1990; Glenberg et al., 1987).

The probability of recall of tokens from a spatial mental model should decrease as the predicted spatial mental model construction order proceeds from first to last. This is a kind of



'negative recency effect'. A similar negative recency effect in long-term memory for word lists has been observed by Craik (1970). Craik interpreted these results as indicating that earlier items were more likely to receive deep, semantic processing (see section 1.6.1). Later items, which are more likely to be produced from primary memory during immediate recall, are less likely to be processed semantically and hence less well retained in long-term memory (Craik & Lockhart, 1972). This explanation is similar to the idea that tokens integrated into a mental model early in the construction process will receive greater elaborative processing. The interpretation here differs from that of Craik in that it is not serial position at presentation (i.e. the order in which items are mentioned in a list or text) which is important, but the order in which the items are integrated into a spatial mental model. Some experimental evidence already exists to support this position. Morton Ann Gernsbacher (1990; 1991) has presented evidence that where discourse supports the construction of a single, coherent mental structure (such as a mental model) items which are mentioned first take longer to process and are accessed more readily. Denhière and Denis (1988) have shown a similar relationship between order of mention and the probability of recall of landmarks in a spatial description. However, a genuine test of this prediction requires that order of mention in the text (text order) and the order of spatial mental model construction are not identical. Only then can the possibility that people recall the language of text be ruled out. The status of order effects in recall as a test of memory for a spatial mental model is discussed in more detail in Chapter 3.

### Memory for consultation processes

Consultation processes in working memory may result in two different kinds of long-term memory trace. Firstly, when a particular spatial relation is accessed by a consultation process the long-term memory trace of that relation is strengthened. This applies both to spatial relations explicit in the original spatial description (consulted during construction) and to new spatial relations that are 'noticed' when the model is consulted at a later stage (Glenberg & Langston, 1992). The idea that memory for relations and tokens in the mental model is dependent on structural processing of this kind is reminiscent of the advantage of relational over item specific imagery in recall (Hunt & Einstein, 1980; Marschark, 1985; Morris & Stevens, 1974). Memory for this kind of consultation process is demonstrated by the retention of the structure or 'gist' of a spatial mental model, but at least one recent account of memory for spatial descriptions has questioned the conclusion that people retain the gist of determinate spatial descriptions in long-term memory (Payne, 1993). The issue of whether people remember the structure (or 'gist') of a spatial mental model is explored in Chapter 4.

The second kind of long-term memory trace that a consultation process might give rise to is memory for a visual image (or particular view) generated from a spatial mental model. This memory trace would be akin to that underlying recognition memory for pictures or scenes. While it is possible that people could remember an image generated from a particular spatial mental



model there is, as yet, little evidence to support this view (Payne, 1993; Tversky, 1991). The argument presented here, however, is that memory for spatial structure does not, in itself, constitute memory for a visual image.

### 2.5.3 Conclusions

Long-term memory for a spatial mental model can take a number of different forms, depending on the particular processes involved in its creation. In order to understand how or whether spatial mental models are remembered it is necessary to consider the processes involved in comprehending a spatial description, the operations involved in its construction and the ways the spatial mental model is consulted. The remaining chapters in this thesis present and discuss evidence from a number of experiments on long-term memory for spatial descriptions with reference to this framework.

## Chapter 3 - Ordinal recall as evidence of memory for a spatial mental model

In Chapter 2 it was argued that a spatial mental model constructed from discourse can be remembered in three relatively distinct ways; memory for the processes of discourse comprehension, memory for the processes of construction and memory for processes used to consult the spatial mental model. In this chapter evidence that people remember an integral aspect of the construction process - the order of spatial mental model construction - is investigated. The experiments presented here concentrate on the findings of Tversky and Taylor (1992a; see also Tversky, 1991) also discussed in Chapter 2.

### 3.1 Challenges to Tversky's account of ordinal effects in drawing recall

The experiments presented in this chapter examine the nature of the drawing order effect observed by Taylor and Tversky (1992a). Tversky has argued that drawing landmarks in the order they were mentioned in a narrative (the drawing order effect) is support for the view that people are remembering the order of spatial mental model construction (Tversky, 1991). There are several problems with this conclusion. The first objection is that the drawing order effect may be an artefact of the particular learning procedures and materials used by Taylor and Tversky (e.g. the stories were presented in segments using a personal computer). The second challenge to Tversky's position is that, while a spatial mental model might preserve the order of construction, no reason is put forward as to *why* this information is preserved. The third and most important challenge is that drawing order may not be determined by memory for a spatial mental model at all. In each of the stories the order of mental model construction corresponds to the order objects are mentioned in the propositions of the text. Thus, the drawing order effect



may not be determined by memory for a spatial mental model, but by memory for the propositions of the text. While construction of a spatial mental model may result in a representation of the language of a text, albeit not a very rich or elaborated one, memory for the language of a text cannot be taken as evidence for the construction of a spatial mental model.

Of the three challenges to Tversky and Taylor raised here the first two are relatively minor. Tversky and Taylor have replicated their initial findings several times and the drawing order effect appears to be fairly robust (Tversky, 1991). However, studies by Kulhavy and colleagues have failed to obtain an effect of order during recall of spatial information (Kulhavy, Stock, Verdi, Rittschof, & Savenye, 1993; Kulhavy, Woodard, Haygood, & Webb, 1993). They interpret their findings as supporting a dual-coding explanation, where the spatial properties of the map are encoded as “intact images in memory” and the narrative is encoded verbally (Kulhavy et al., 1993). In contrast, Tversky has argued that the spatial representation people form on these tasks is not image-like (Tversky, 1991).<sup>1</sup> The Kulhavy studies used learning procedures where subjects were given a map to learn in addition to being presented with a text describing the environment depicted by the map. This means that the order of presentation of landmarks in the text and the order landmarks are scanned in the map differ. Therefore it is not surprising that the eventual order and probability of recall showed no clear relationship with the order landmarks were encoded during presentation of either the map or the text. In one study (Kulhavy et al., 1993) verbal protocols were produced by subjects while they learned the maps. Before reading a text about the environment depicted in the map the subjects were asked to reconstruct the map. No systematic relationship was observed between serial position of the landmark in the verbal protocol and the accuracy of its spatial location or probability of recall in the reconstruction of the map. However, the *order* of reconstruction is not reported, possibly because only a few subjects recalled all the features depicted in the map. The authors also note that no marked primacy or recency effects were present. It appears, therefore the drawing order effect obtained by Taylor and Tversky (1992a) is dependent on a single consistent order of presentation. If order of presentation varies between subjects then the predicted order of spatial mental model construction will also vary from subject to subject. It is also possible that recall during map reconstruction is not influenced by the order in which features are encoded during the learning of a map. These findings can therefore not be considered as conclusive evidence against the position adopted by Tversky. In addition, the work of Payne (1993) provides some support for Tversky’s position. It suggests that memory for the order of spatial mental model construction could be mediated by an *episodic construction trace* (see section 2.4.5).

---

<sup>1</sup> It should be noted that both Kulhavy and Tversky stress the importance of the structural properties of the representation. However, Kulhavy argues that the representation is an intact “unitary chunk” in long-term memory whereas Tversky believes that the representation is composed of parts and relations among parts (e.g. as a structural description).

However, the final objection is far more serious. Tversky provides no direct evidence that the order effect is not mediated by memory for the language of the text that has been read. Perhaps, readers were simply remembering the text and not a spatial mental model constructed from the text?

This chapter examines the nature of the drawing order effect observed by Taylor and Tversky. Three experiments are presented which focus on the relationships between text order, spatial mental model construction order and the order of recall in a drawing task. For this reason subjects in the experiment are asked to read stories for which the text order and the spatial mental model construction order differ. This is accomplished by presenting some readers with stories that depart from their natural or original sequences - in these 'scrambled' stories sections of the text are presented in a random order (Kintsch, Mandel, & Kozminsky, 1977). This manipulation disrupts the original text order. However, in order to have a coherent spatial representation of the situation described by the text it is necessary to construct a spatial mental model like that of the original (unscrambled) order of the story. A second reason for the use of 'scrambled' narratives in these experiments is that some researchers have argued that being able to understand a scrambled story requires the construction of a mental model (van Dijk & Kintsch, 1983).

### 3.2 Preface to the experimentation

In the nine experiments reported in this thesis distribution-free are preferred to parametric statistics for the majority of analyses. There are two principle reasons for this decision. The first reason is that many of the analyses are for a predicted trend or order (e.g. the order in which landmarks are recalled in a drawing task). It is not necessarily the case that these predicted trends are linear therefore a more conservative assumption (that they are monotonic) can be tested by applying one of the three distribution-free trend tests adopted here (Meddis, 1984; Siegel & Castellan, 1988). In the specific case of the analysis of drawing orders it also questionable whether the data are interval, therefore statistics based on ranking procedures tests (and which only assume ordinal data) seem more appropriate. The second reason concerns the occurrence of ceiling and floor effects in some of the data sets reported in the thesis. Where ceiling or floor effects are present parametric data may longer be normally distributed (and variances for different conditions may not be homogenous). Again, in these cases, distribution-free statistics are preferred.



### 3.3 Experiment 1

The first experiment attempted to replicate Taylor and Tversky's drawing order effect with both unscrambled and scrambled texts. Subjects were presented with either a *route* or a *survey* text in either an *unscrambled* or *scrambled* format. This resulted in a total of four texts. Both the route and the survey text described a small village. The route text described a linear, coreferentially coherent text and took the perspective of a character driving round the village in a specific sequence. It used a mixture of canonical and relative directions. The survey text was coherent but did not describe the village in a linear fashion; the perspective it took was that of a character describing the village as if describing a plan (see sections 1.2.2 and 2.4.6 for a discussion of route and survey representations). It used only canonical directions (for a copy of the unscrambled survey and route text refer to Figures 3.1a and 3.1b respectively). Previous research has shown that scrambled *survey* texts can be understood even when it is not possible for subjects to determine the original order the sentences were in (Langer, Keenan, & Nelson, 1991; Langer, Keenan, & Nelson, 1991). This appears to be because survey texts rely more on global, canonical spatial relations (e.g. North or South). Route texts rely more on local, relative spatial relations (e.g. right and left). Global, canonical spatial relations allow people to understand a scrambled text without establishing the original order of the narrative.

Tversky's explanation of the drawing order effect would predict that drawing order should always be influenced by the original *unscrambled* order of the text that was read. This should happen whether or not the sentences are actually presented in that unscrambled order. However, for a drawing order effect to be apparent readers would have to remember a significant proportion of the landmarks from the text version they had read. Where not all the landmarks are recalled the spatial mental model account suggests that items integrated into the spatial mental model early in the construction process will be remembered better than those entered later on. Some support for this prediction is found in a study by Denhière and Denis (1988). They gave readers texts describing an island and then asked them to fill in blank maps of the island. They found that the order of a landmark in the text, but not the time taken to read sentences, predicted the probability of recall.

The following four predictions are made. Firstly, for the *unscrambled* text conditions the drawing orders of subjects should reflect the order of spatial mental model construction. Secondly, for the *scrambled* texts drawing orders should be closer to the spatial mental model construction order than to the order landmarks are actually mentioned in the scrambled text (text order). However, as has already been noted, subjects may be able to understand the *scrambled survey* text without attempting to identify the order of the original unscrambled survey text. If this happens drawing order for the scrambled survey text should reflect the order landmarks are actually mentioned in the scrambled survey text. If drawing recall is not at (or close to) ceiling, it

may not be possible to detect any drawing order effects. Accordingly, the third prediction is that the probability of recalling a landmark should also be influenced by the order of spatial mental model construction rather than by text order. The fourth and final prediction is that reading times should also be influenced by the scrambling manipulation, reflecting the increased difficulty of constructing a spatial mental model from a scrambled text. This is because the text would probably have to be read several times before people are able to construct a spatial mental model.

### 3.3.1 Method

#### Subjects

The forty-seven subjects in the experiment were graduate students or staff from the Open University. Ages of subjects ranged from early twenties to early sixties.

#### Design

There were two between subject variables. The first independent variable was the spatial perspective adopted by each story (route or survey). The second independent variable was text organization; the original or natural order of the story (henceforth unscrambled) and a randomly ordered (or scrambled) version. Subjects were randomly allocated to one of the four text conditions.

#### Materials

Two versions of a story entitled “Eric goes to Lower Barking” were prepared. Both versions described an inexperienced detective attempting to locate a car in a village called Lower Barking. The key differences between the texts lay in the form of spatial description used. The first text took a survey perspective (see Figure 3.1a). It described the village as seen from above, using canonical direction terms such as ‘North’, ‘South’, ‘East’ or ‘West’. The second text took a route perspective as if travelling round the village in a particular sequence (see Figure 3.1b). The route text used relative directional terms such as ‘left’, ‘right’ and ‘clockwise’ (however, in order to reduce confusion and increase coherence for the scrambled version, some canonical terms were also used). Additional non-spatial information was added to the survey text in order to make the two texts approximately equal in length (both were eighteen sentences and approximately 285 words long). The two texts were also comparable on a selection of text comprehension measures. Scrambled versions of the texts were obtained by grouping consecutive sentences into pairs and randomly reordering the resulting sentence pairs. In this experiment sentence pairs were used because subjects found that reordering by sentence was too difficult when the materials were



piloted. The same random number sequence was used to scramble for the survey and route versions of the text in order to equate the degree of 'disarray' present in each scrambled text.

Eric goes to Lower Barking (unscrambled survey text).

Eric was the new man in the Milton Keynes Detective Agency and his first case was a routine car surveillance job. Eric followed the suspect's red Ford Escort south down the motorway until the suspect turned off.

Eric ended up in the picturesque village of Lower Barking which lay about two miles east of the motorway. Eric couldn't help admiring the unspoilt rural landscape.

The *Main Road* (1) had taken Eric straight into the centre of the sleepy hamlet. Suddenly Eric realised that he had lost sight of the car he was supposed to be following.

The inexperienced detective would have to drive round the little village until he tracked down his quarry. *The Common*, (2) in the centre of the village, was circled by Dog Kennel Lane.

*Old Farm Lane* (3) ran due north from the Common until it came to a *Deserted Farmhouse* (4). *Church Lane* (5) ran east from the Common, past a red telephone box, until it reached *St. Malcolm's* (6).

Eric had arrived in the village on the Main Road from the West. Lower Barking began where the Main Road met Dog Kennel Lane in a T-junction.

On the northern corner of the Main Road and Dog Kennel Lane stood the village *Post Office* (7). Directly south of the Post Office Eric noticed a pub car park and just beyond it the *Rose and Crown Inn* (8).

After driving round the village for forty minutes Eric decided to end his search. The disappointed detective made his way to the Rose and Crown for a consolatory drink.

As Eric got out of his car he brushed against a nearby vehicle. The car, a red Ford Escort, began to emit a high pitched wail ...

Figure 3.1a: The unscrambled survey text presented in Experiment 1 (landmarks used in the analysis are shown in italics and numbered in the order in which they are first mentioned).

Eric goes to Lower Barking (unscrambled route text).

Eric was the new man in the Milton Keynes Detective Agency. His first case was a routine surveillance job.

Eric followed the suspect's red Ford Escort south down the motorway until he turned off towards Lower Barking which lay about two miles east of the motorway along a *Main Road* (1). The Main Road took Eric right into the picturesque little village.

Suddenly Eric realised that he had lost sight of the suspect's car. Eric would have to explore the village to find his quarry.

Driving into the village from the West, Eric slowed down where the Main Road ended in a T-junction with *Dog Kennel Lane* (2). To Eric's left, on the corner of the Main Road and Dog Kennel Lane, stood the village *Post Office* (3).

To his right Eric noticed a car park and beyond it the *Rose and Crown Inn* (4). Eric turned left onto the northern branch of Dog Kennel Lane which curved round to the right and circled the village Common.

Next Eric turned left, due North up *Old Farm Lane* (5), until reached a *Deserted Farmhouse* (6). Eric turned round and drove back to the Common, continuing clockwise round Dog Kennel Lane until he came to a junction where a left turn took him onto *Church Lane* (7).

Driving East down Church Lane Eric went past a red telephone box, until he reached *St. Malcolm's* (8). Eric had driven round the village for forty minutes and he decided to give up the search.

From St. Malcolm's Eric drove back to Dog Kennel Lane and continued clockwise round the Common until he returned to the junction with the Main Road. The disappointed detective drove into the car park of the Rose and Crown for a consolatory drink.

As Eric got out of his car he brushed against a nearby vehicle. The car, a red Ford Escort, began to emit a high pitched wail ...

Figure 3.1b: The unscrambled route text presented in Experiment 1 (landmarks used in the analysis are shown in italics and numbered in the order in which they are first mentioned).

### Procedure

Each subject was informed that they would be given a short story to read, followed by a task related to the events described in the story. They were instructed to read the story through as many times as they wished. They were asked to concentrate on the spatial layout and physical

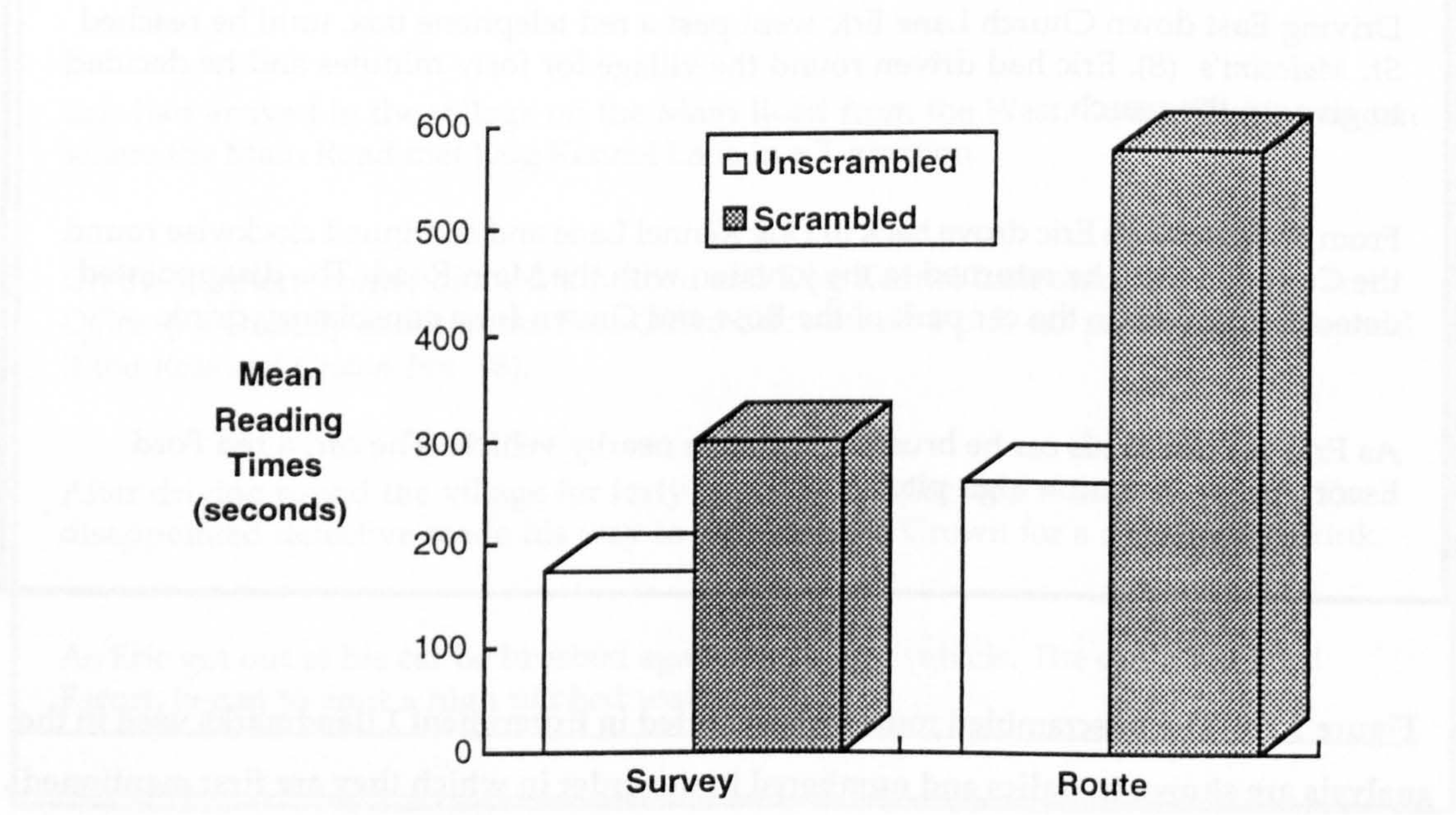


description of the village mentioned in the story. When they had finished reading and had a clear understanding of the events in the story they were asked to indicate this to the experimenter. Reading times were recorded for each subject. Those receiving scrambled versions of the texts were informed that the story they were about to read contained sentences out of their original or natural sequence. After reading the story each subject was asked to draw the major features (such as roads and buildings) that make up the village of Lower Barking as well as they could. In addition each subject was asked to label features and landmarks as they drew them. At the same time as the subject drew the village the experimenter noted down the order in which each landmark was drawn. To make scoring easier a coding sheet, consisting of a table containing every possible combination of landmark and output position, was constructed. Recording the drawing order was thus a simple matter of ticking the appropriate box.

### 3.3.2 Results

#### Reading times

Reading times were analyzed using a 2 by 2 factorial ANOVA. There was a significant main effect of text perspective; survey versus route ( $F_{1,43} = 10.80$ ;  $p < 0.01$ ) and text organization; unscrambled versus scrambled ( $F_{1,43} = 16.23$ ;  $p < 0.001$ ). Mean reading times, in seconds, for the four conditions are given in **Figure 3.2** below:



**Figure 3.2: Mean reading times by text perspective and text organization**



There was a trend towards an interaction between text organization and perspective, but this did not reach significance ( $F_{1,43} = 2.88$ ;  $p = 0.097$ ). Reading and understanding a scrambled text takes longer than an unscrambled text. This is consistent with the prediction that people reading a scrambled text have greater difficulty in constructing a spatial mental model of the situation the text describes, perhaps because of factors such as a reduction in referential continuity which are known to make spatial mental model construction more difficult (Ehrlich & Johnson-Laird, 1982).

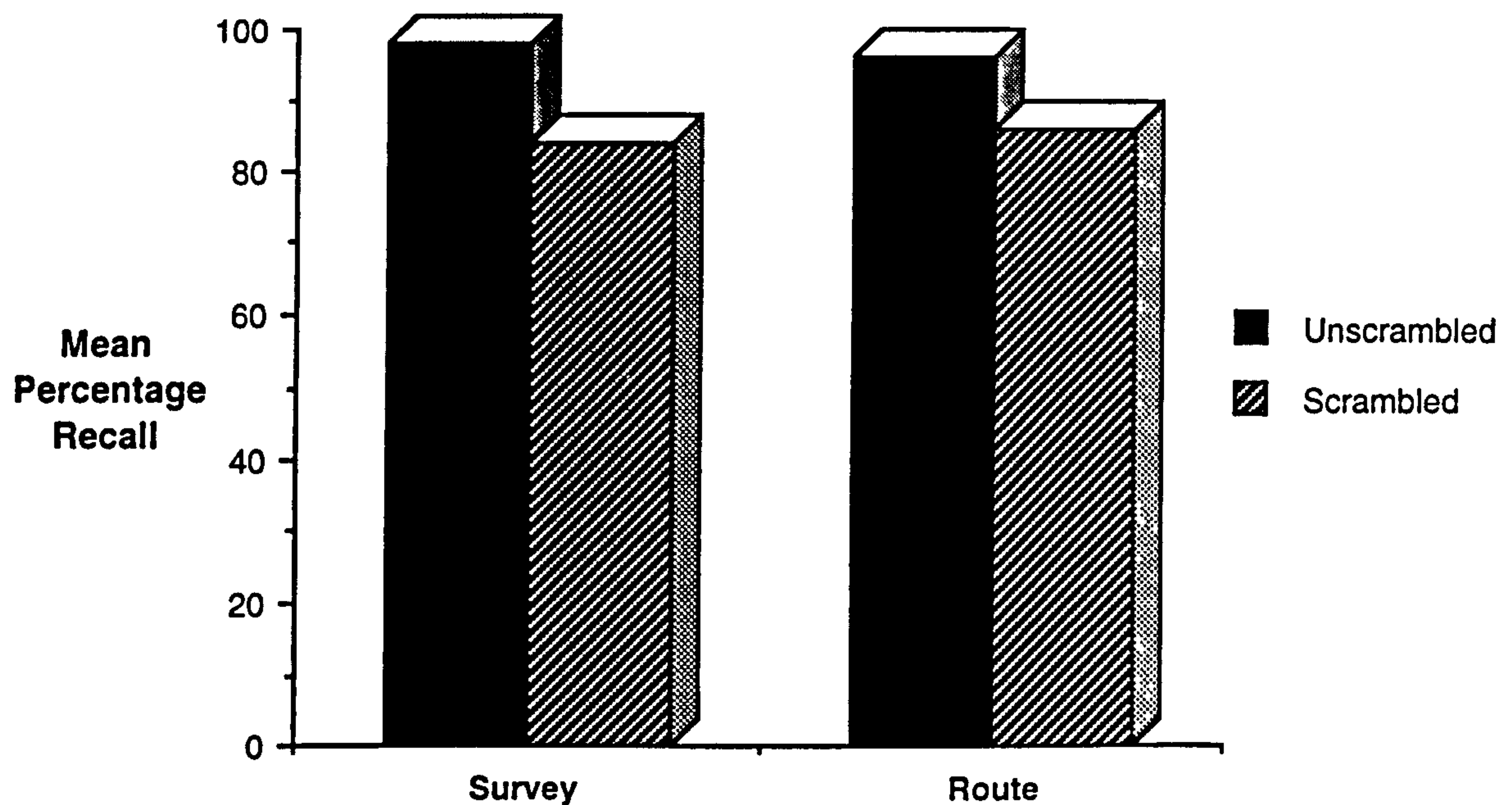
### **Scoring procedures for the drawing task**

Preliminary analysis of the maps drawn by subjects revealed that very few subjects recalled all of the landmarks and features mentioned in the story. This appeared to be due to two factors. Firstly, not all subjects interpreted the instructions provided by the experimenter in the same way (refer to the procedure section for the instructions which were provided). Some subjects later reported they had (not unreasonably) excluded landmarks such as 'Milton Keynes' which were not "in the village". Other subjects reported that they had not included features such as the 'red telephone box' because they were not landmarks (i.e. roads and buildings). The second factor was that many subjects, particularly those in the scrambled text conditions, were unable to remember all the landmarks and features present in the story. For this reason the analysis was confined only to the major landmarks which were within or partly within the village (these landmarks are shown in italics in Figures 3.1a and 3.1b). In this way it was possible to reduce any differences due to how subjects interpreted the instructions for the drawing order task. One other minor problem in scoring recall on the drawing task should also be noted. Distinguishing between 'The Common' and the road 'Dog Kennel Lane' (which surrounds it) in any given drawing proved to be very difficult and in many cases impossible. For this reason the two landmarks were considered as a single item for the purposes of data analysis. It is extremely unlikely this had a significant impact on the results as, if both landmarks were recalled, they were (without exception) recalled together.

### **Levels of recall on the drawing task**

Almost all the subjects in the unscrambled conditions recalled all eight of the major landmarks. This ceiling effect made it difficult to investigate differences in recall between groups. Between group differences were analyzed using the Mann-Whitney U test. Percentage mean recall of these landmarks is shown by group in Figure 3.3 below:





**Figure 3.3: Mean recall by text perspective and text organization.**

There was a significant effect of scrambling, but no difference between the route and survey perspective (though it is possible that ceiling effects may have obscured an interaction). Scrambled text versions resulted in significantly lower recall than unscrambled texts regardless of perspective ( $U = 183$ ;  $N = 47$ ,  $p = 0.018$ ).<sup>2</sup> This is consistent with the view that unscrambled texts result in superior recall because it easier to construct a single, coherent spatial mental model from them. Again, however, this effect could also be caused by linguistic factors (such as loss of referential continuity in the scrambled text).

### Drawing orders

Drawing orders were obtained for those maps which depicted all eight major landmarks (refer to the procedure section for a description of how drawing orders were recorded). For each of the four texts an order of landmarks was obtained which corresponded to the order in which items were mentioned in the text (henceforth referred to as *text order*). By comparing *drawing order* with text order it is possible to calculate a measure of the extent to which the two orders are related. The measure adopted was the Kendall rank-order correlation coefficient or *tau* statistic. Tau was chosen because it is calculated by counting the number of times individual scores have to be swapped (inverted) in order to transform the ranks of one data set into those of another data set. So tau is used here not as a correlation coefficient but a measure of how closely two orders are related. An additional advantage is that tau can be used to 'partial out' the influence of a third

---

<sup>2</sup> This analysis is still significant if all thirteen landmarks, rather than just the eight major landmarks are included in the analysis.

relationship, provide that relationship is known to exist (Howell, 1992; Siegel & Castellan, 1988).

It was also possible to calculate a tau value for how each drawing order differed from the text order of each of the three other texts used in the experiment (i.e. the text orders of the three texts the subjects had *not* read). Thus it was possible, for each subject, to calculate a tau value for the relationship between their drawing order and *any* of the four text orders. Using these tau values it is possible to determine which text order was most closely related to the drawing orders produced by subjects who had only one of the four possible texts. This was carried out by comparing tau values for a given drawing and text order for each group of subject in turn. These comparisons were made using within subjects t-tests.

Condition	Tau values compared between drawing order and text order 1 or 2.	
	1. The text read by the subject	2. Comparison text (not read by the subject)
US	Unscrambled survey text	Unscrambled <i>route</i> text
UR	Unscrambled route text	Unscrambled <i>survey</i> text
SS	Scrambled survey text	<i>Unscrambled</i> survey text
SR	Scrambled route text	<i>Unscrambled</i> route text

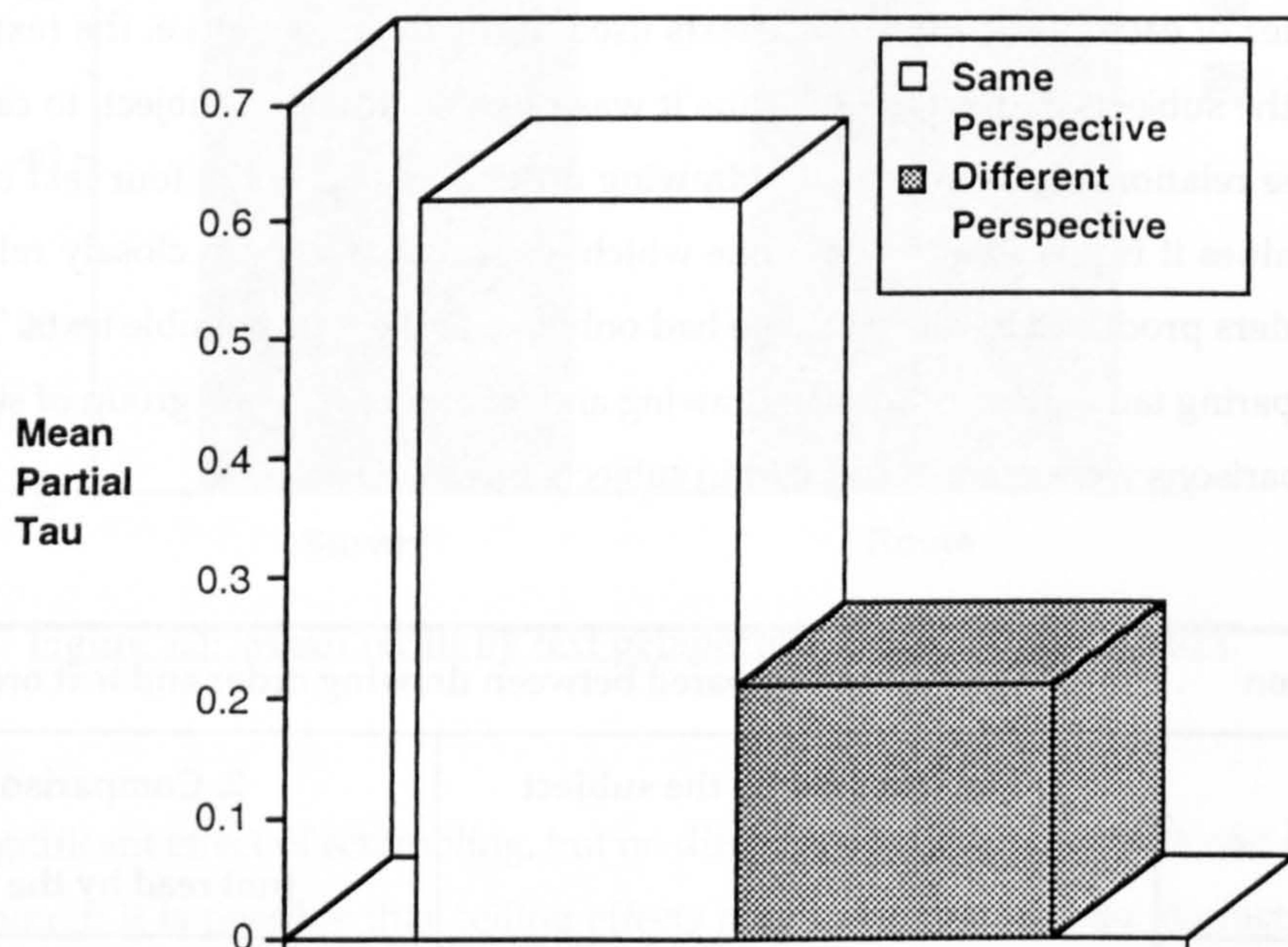
**Figure 3.4: Table of the choice of comparison text in the drawing order analysis for each condition.**

Only a selection of the possible comparisons were conducted. In each case the tau value between drawing order and the text order of the text the subjects had read was compared with that between drawing order and text order for a text the subject had *not* read. These comparison texts are listed in Figure 3.4. For subjects in the unscrambled text conditions the comparison text was the unscrambled text which the subject had not read. The comparison text (a text which described the village from a different perspective) provides a control or ‘baseline’ tau value. This control tau value would take into account any general preferences on the part of the subjects for a particular drawing order (e.g. right to left or top to bottom). For these two comparisons partial tau values were used (Siegel & Castellan, 1988). These partial tau values take account of the positive relationship between the text order of the two unscrambled texts (for example both text orders begin with the ‘Main Road’).

The drawing orders obtained from readers of the unscrambled texts were significantly closer to the order items were mentioned in the same perspective text (the text that they had



read) than that of a different perspective text. This was confirmed by carrying out related t-tests for the unscrambled survey ( $t= 3.33$ ; d.f. = 8, one-tailed  $p = 0.0051$ ) and unscrambled route conditions ( $t= 3.28$ ; d.f. = 8, one-tailed  $p = 0.0056$ ). This difference was significant even when partial tau values were not used. Mean partial tau values for same and different perspectives are shown in Figure 3.5 below:



**Figure 3.5: Mean concordance values by perspective for the unscrambled texts**

This finding replicates the results obtained by Taylor and Tversky (1992a). It suggests that drawing orders on this task may be related to the order of spatial mental model construction. However, in this comparison text order and predicted spatial mental model construction order are confounded.

For subjects in the two scrambled text conditions the comparison text was the unscrambled text which adopted the same perspective as the scrambled text they had read. Thus the comparison was between the text order of the scrambled text they *had* read and the predicted order of spatial mental model construction. This is because constructing a spatial mental model construction for a scrambled text should require readers to reconstruct the situation described by the original unscrambled text. So people should construct a spatial mental model from a scrambled text in an order which is as close as possible to the order of the corresponding unscrambled text.

For the readers of the scrambled texts there were only small differences between the tau values generated for the text order of the scrambled texts and those of the comparison text. These differences were not significant. This was confirmed using related t-tests for both the scrambled survey condition (mean SS tau = 0.53, mean US tau = 0.53,  $t= 0.002$ ; d.f. = 8, two-tailed  $p = n.s.$ ) and the scrambled route condition (mean SR tau = 0.44, mean UR tau = 0.42,  $t= -0.13$ ; d.f. = 6, two-



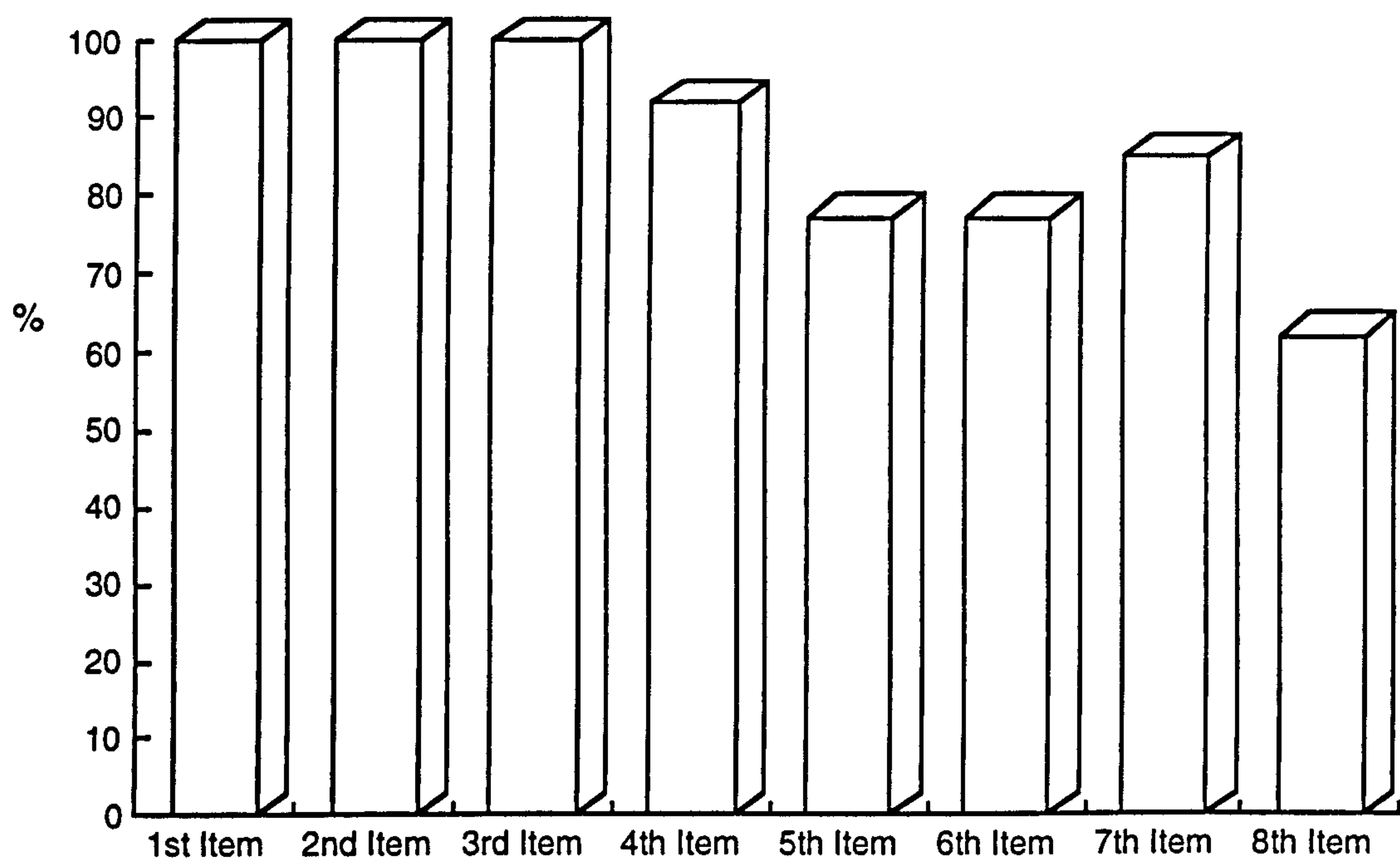
tailed  $p = 0.93$ ). The lack of significant differences is likely to be due to three factors. Firstly, many subjects in the two scrambled conditions could not be included in the analysis because they recalled less than eight of the major landmarks. Secondly, drawing orders for the *scrambled survey* condition were quite strongly related to the order of the unscrambled route (though this trend did not reach significance). This may have been influenced by a general tendency to draw from left to right (approximately the order in which items are mentioned in the unscrambled route text). Thirdly, the scrambling procedures resulted in relatively high residual correlations between the different text orders. These correlations were calculated using Kendall's tau. The residual correlations for the unscrambled survey text was 0.327 (with the unscrambled route text), 0.357 (with the scrambled survey text) and 0.429 (with the scrambled route text). The residual correlations for the unscrambled route text was 0.837 (with the scrambled survey text) and 0.546 (with the scrambled route text). There was also a high residual correlation between the scrambled route and survey text orders ( $\tau = 0.5$ ).

### Probability of landmark recall (scrambled texts conditions only)

Because many of the subjects in the scrambled text conditions failed to recall all eight major landmarks further analyses were carried out on the probability of recall. Recall of items from the unscrambled stories was at ceiling and so was not analyzed. For drawings made by subjects in the scrambled text conditions the presence or absence of each of the eight major landmarks in the story was recorded. Then the likelihood of recall was compared within subjects against one of two predicted orders; either the order of the unscrambled text (the predicted spatial mental model construction order) or the order of the scrambled text (the actual order of the presented text). The data were analyzed using a non-parametric trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). Each set of data was tested (using the set of coefficients: -7, -5, -3, -1, 1, 3, 5, 7) for a trend against the same perspective scrambled and unscrambled text order in turn.

For the route text no significant relationship was found between probability of recall and the actual text order read by the subjects (the order of the scrambled route text). Instead, probability of recall appeared to be related to the order of the unscrambled route text (the predicted order of spatial mental model construction). This was confirmed by testing for a specific trend using the trend test for dichotomous data described above ( $Z$  estimate = 3.92,  $N = 14$ , one-tailed  $p < 0.0001$ ). Figure 3.6 shows the mean percentage recall by item for readers of the *scrambled* route text plotted against the order of mention of items in the *unscrambled* route text:

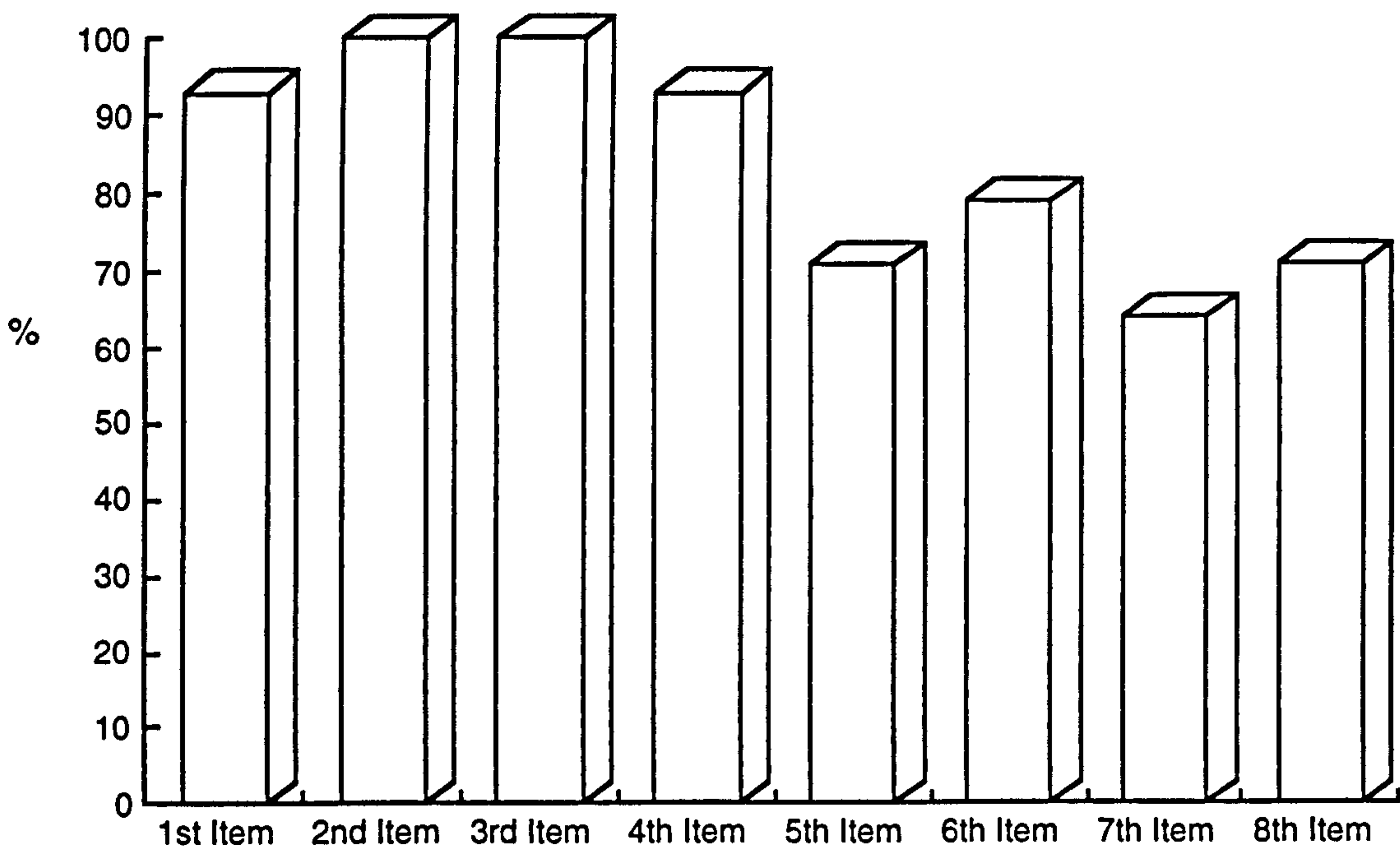




**Figure 3.6: Mean percentage recall for readers of the scrambled route text against order of mention in the unscrambled route text**

This result is clear support for the prediction that the probability that a landmark is recalled from a scrambled route text is related to the order of spatial mental model construction.

For the scrambled survey text the results were reversed. There was no significant relationship between the probability of recall and the order of the unscrambled text (the predicted order of spatial mental model construction).



**Figure 3.7: Mean percentage recall for readers of the scrambled survey text against order of mention in the scrambled survey text**

Instead probability of recall was related to the order of the text subjects had read (the order of the scrambled survey text). Significance was confirmed by testing for a specific trend ( $Z$  estimate = 3.89,  $N = 13$ , one-tailed  $p < 0.0001$ ). Figure 3.7 shows the mean percentage recall by item for readers of the scrambled survey text plotted against the order of mention of items in the text they had read. This result suggests that readers of scrambled survey tests do not construct a spatial mental model in the order of the unscrambled survey text. Two possible interpretations of this result are offered. The first explanation is that people are not recalling a spatial mental model, but a propositional representation of the language of the text. The second explanation is that people are able to construct a spatial mental model from a scrambled survey text without identifying the order of the unscrambled text. This explanation is supported by experiments which show that scrambled survey texts can be understood and remembered even when the original order of the text can not be detected (Langer, Keenan, & Cumbo, 1992; Langer et al., 1991). This probably arises because the global, canonical direction terms used in survey texts are easier to understand when a text is scrambled than local, relative direction terms.

### 3.3.3 Discussion

In the introduction to this experiment it was predicted that drawing order and the probability of recall on a drawing task should reflect the order of spatial mental model construction. For the two unscrambled texts this is the order landmarks are mentioned in the text subjects have read. For the scrambled text conditions, however, this is the order landmarks are mentioned in the corresponding unscrambled text (a text which the subjects have not read). This is because people will try to construct a spatial mental model based on the situation described by the text. This should involve trying to construct a spatial mental model as similar as possible to that supported by the unscrambled version of the text (the unscrambled or 'natural' order of the story).

The first important finding presented here is that the drawing order effect observed by Taylor and Tversky (1992a) for *unscrambled* texts has been replicated using different materials and different learning procedures. However, for the *scrambled* texts no significant relationship has been obtained between drawing order and either the order landmarks are mentioned in the text or the predicted order of spatial mental model construction. The reading times for the four text versions replicate the common finding that scrambled stories require longer to read than unscrambled stories (Langer et al., 1992; Langer et al., 1991).

Unfortunately, many readers appeared not to have constructed a coherent spatial mental model (or at least not a complete one). This may be reflected in the failure to find any significant differences between tau values for the text orders of the scrambled texts and relevant



unscrambled text order (see Figure 3.3). This failure may be due to the fact that half the subjects in the scrambled route condition were not included in the drawing order analysis, because they could not recall the major landmarks in the village. Another factor may have been a preference among some subjects to adopt a left to right drawing order which, by coincidence, was similar to that of the unscrambled route text. Similar preferences have been observed by other researchers when people are asked to draw maps of places they know or have learned about (Taylor & Tversky, 1992a).

However, some support for the hypothesis that readers of scrambled texts attempt to construct a spatial mental model close to that of the corresponding unscrambled text is provided by evidence concerning the probability of recalling a landmark. Readers of the scrambled route text showed a significant relationship between probability of landmark recall and the predicted order of spatial mental model construction. People who read the scrambled route text were more likely to remember landmarks which were mentioned early in the corresponding unscrambled text than those mentioned late. This suggests that readers of the scrambled route text attempted to construct a spatial mental model of the situation described by the unscrambled route text and that they were at least partly successful. However, a completely different result was obtained with the scrambled survey text. Readers of the scrambled survey text showed no relationship between the probability of recalling a landmark and the predicted order of spatial mental model construction. In fact, the probability of recalling a landmark from the scrambled survey text was significantly related to the order landmarks were mentioned in the scrambled survey text itself. This is consistent with either of two different explanations. The first explanation is that readers of the scrambled survey text were unable to construct a spatial mental model. In this case drawing recall would be based on a propositional representation of the language of the text. The second explanation is that the global, canonical direction terms used in the scrambled survey text allow subjects to construct a spatial mental model without having to identify the original (unscrambled) order of the survey text. In this case the actual and predicted order of spatial mental model construction will be completely different. Evidence to support this explanation comes from studies by Langer which have shown that scrambled survey texts can be very well understood and remembered even though people are very poor at reordering them (Langer, Keenan, & Bergman, 1993; Langer et al., 1992; Langer et al., 1991).

In conclusion, the evidence from this experiment provides some support for the view that the order and probability of landmark recall on a drawing task is predicted by the order of spatial mental model construction (Denhière & Denis, 1988; Gernsbacher, 1991; Tversky, 1991). However, the drawing recall data from readers of the scrambled survey text could also be consistent with a linguistic representation of the text. These conclusions are further weakened by several methodological problems. Firstly, unless recall is at ceiling it is not possible to analyze the drawing orders obtained from every subject. Secondly, there is some evidence that drawing orders are influenced by other factors (notably a preference for a left to right drawing order).

Thirdly, and most importantly, it is possible that the instructions given to subjects may have biased them to report certain landmarks more than others. In Experiments 2 and 3 steps are taken to control for all three of these problems.



## 3.4 Experiment 2

This experiment attempts to investigate further the drawing orders produced by readers of scrambled texts. From Experiment 1 there is some evidence that the *probability of recall* in scrambled stories is related to the order of spatial mental model construction. It is likely that the failure to find a relationship between *drawing order* and either the order of spatial mental model construction or the order of the scrambled texts in Experiment 1 was, in part, due to the low levels of recall for the scrambled stories. There were also problems with the methodology and materials adopted in Experiment 1. For this reason Experiment 2 used more coherent stories, which maintained referential coherence from one sentence to the next, which described events in a plausible sequence and which contained some redundant spatial information (typically in the form of canonical directions). These changes were designed to produce a higher level of recall for the two texts used in the experiment. Care was taken so that neither text described a simple left to right progression. On this occasion two route texts were used. Survey texts were not used for two reasons. The first reason is that it is more difficult to construct a plausible, coherent survey text than a plausible, coherent route text. The second reason is that it may not be possible to distinguish between a propositional representation of a text and a spatial mental model constructed from a survey text (refer to the discussion section of Experiment 1). In addition, the study used scrambled texts which were more sensitive to drawing order effects by manipulating the distance between the spatial mental model construction order and the text order of the scrambled text (these changes are discussed in more detail below). Finally, in the instructions for the drawing task subjects were clearly asked to draw *all* the landmarks and features they could remember from the original text. These changes in instructions and materials were designed to address the methodological problems which arose in Experiment 1. The two stories used in this experiment are shown in Figures 3.8a and 3.8b. The title and first sentence set the scene for each story and are not included in the scrambling procedure. Italics indicate the ten landmarks for which drawing orders are recorded. Numbers in brackets indicate the serial position of each landmark in the unscrambled text.

Before being asked to draw the environments they had learned each reader was also tested on their recognition memory for the sentences of the texts they had read. This involved presenting a series of sentences by computer and asking subjects to decide as quickly and accurately as possible whether each sentence was true or false of the story they had read. Four types of sentence were used: verbatim sentences (identical to sentences in the original text), paraphrase sentences (rewordings of the verbatim sentences), inference sentences (new sentences which were true with respect to the situation described by the original text) and distracter items (which contradicted sentences contained in the original text).

Roger Tenement Goes Shopping (unscrambled shop text).

Every Tuesday Roger Tenement visits the Music Store in Central Milton Keynes.

The *Main Entrance* (1) to the Music Store is due North of the *Bus Stop* (2) where Roger arrives in town.

As Roger enters the shop through the sliding doors he notices a new *Video display* (3) to his left.

Behind the Video display Roger can see the familiar red sign of the *Bargain Section* (4) of the shop and he goes over to inspect the special offers.

After searching through the bargain section, Roger walks down the Western wall to look at the *Rock Music Section* (5) at the back of the shop.

While browsing through the Rock Music section Roger realizes that the *Jazz Section* (6) has been moved to the back of the shop and is on his right.

Roger walks over to the Jazz section and spots an interesting CD which he takes to the *Information Counter* (7) at the front of the shop in the South East corner.

At the information counter Roger asks if he can listen to the CD and he is directed to a *Listening Booth* (8) in the centre of the shop.

After listening to his CD Roger leaves the booth and pays for the CD at the *Cash Desk* (9) on the Eastern wall of the shop.

With the newly purchased CD in his hand Roger leaves through the *Side Entrance* (10) next to the cash desk.

Figure 3.8a: The unscrambled version of the shop text.

The sentences in the recognition test were presented to each subject in a random order generated by the computer. After the recognition tests the readers were given an opportunity to reread each story once. The purpose of this second presentation of the original texts was to prevent any contamination from the recognition memory test. Each subject would be able to reread the original text and check whether their responses in the recognition test were accurate or not.



A Morning in the Life of Erica Emerald (unscrambled apartment text).

Erica Emerald lives alone in a two room studio flat on the outskirts of Milton Keynes.

Every morning at five o'clock precisely Erica wakes up in the *King-size Bed* (1) in the centre of her *Plush Bedroom* (2).

After getting up Erica makes her way over to the *Ensuite Bathroom* (3) which lies against the middle of the North facing wall of the bedroom.

When Erica has finished her morning shower she takes some clothes from the *Large Wardrobe* (4) positioned in the centre of the wall, opposite the ensuite bathroom.

After getting dressed and closing the wardrobe Erica goes through the *Pine Door* (5) located in the wall immediately to the East of the wardrobe.

Having left the bedroom Erica turns right and walks down the short, *Well-lit Hallway* (6).

At the end of the hallway Erica comes to an *Ornate Glass Door* (7) which she opens and walks through.

The glass door leads directly into Erica's *Well-furnished Living Room* (8).

After entering the living room Erica turns to her right and walks down to the far end of the room until she reaches the *Kitchen Area* (9).

When Erica has finished her breakfast in the small kitchen she goes over to the South side of the room and leaves her flat by the *Back Door* (10) located in the middle of the wall.

Figure 3.8b: The unscrambled version of the apartment text.

Each subject was given one unscrambled and one scrambled route story to read. The texts were scrambled by sentence rather than by sections (this was to reduce any relationship which might exist between the predicted spatial mental model construction order and the scrambled text orders). Only route texts were used because readers appear to be better at unscrambling them. One of four different scrambled story orders were used for each subject (this made it unlikely that the results were influenced by a particular scrambled order). Care was taken to make sure that the stories described routes that were not a simple left to right progression - an order which many people seem to adopt when asked to draw maps of places (Taylor & Tversky, 1992a). Finally, each of the original (unscrambled) stories was constructed so that the order of mention in

the text and the predicted order of spatial mental model construction differed. This was done by manipulating the sentence that introduced the first two landmarks (e.g. “The main entrance to the Music Store is due North of the bus stop where Roger arrives in town”). Because the initial landmark in the sentence (the ‘main entrance’) is described in relation to the second landmark (‘the bus stop’) the spatial mental model account predicts that the reference point (‘the bus stop’) is entered into the model first.

The spatial mental model account predicts that drawing order should be significantly related to spatial mental model construction order for both scrambled and unscrambled route narratives. In addition, drawing order for the scrambled texts should be significantly closer to the order of spatial mental model construction than that of the actual scrambled text. It is also predicted that accuracy and response time to the different statements in the sentence verification task should follow the pattern found in experiments by Tversky and others (Perrig & Kintsch, 1985; Taylor & Tversky, 1992a; Tversky, 1991). Verbatim statement should be faster and more accurately responded to than paraphrase sentences which, in turn, should be faster and more accurate than inference statements. This pattern would be in accord with the view that readers form a propositional representation of the text which preserves some information about the verbatim content of the original text.

### 3.4.1 Method

#### Subjects

Sixteen graduate students and members of staff from the Open University participated in the study. The age of the subjects ranged from early twenties to early fifties.

#### Design

There were two within subject variables; text organization (unscrambled and scrambled) and the location described by the text (an apartment and a shop). Half the subjects received the shop text in its unscrambled form and the apartment text in its scrambled form. The remaining subjects received the shop text in a scrambled form and the apartment text in an unscrambled form. The order of the scrambled narrative (one of four random sequences of sentences) varied between subjects. Text organization and the location described by the text were counterbalanced. Subjects were assigned randomly to one of the counterbalanced conditions and to one of the four scrambled text orders.



## Materials

The two stories used in the experiment “Roger Tenement Goes Shopping” and “A Morning in the Life of Erica Emerald” are presented in Figures 3.8a and 3.8b. Both unscrambled texts were ten sentences long (228 and 218 words respectively). The initial sentence after the title was used to introduce the description, but was not scrambled in any of the text versions. The initial sentence did not introduce any of the items involved in the subsequent analysis. The second sentence introduced two items; the first item being described in relation to the second (“The main entrance to the Music Store is due North of the bus stop where Roger arrives in town.”). Each of the additional eight sentences introduced one new item and described its location relative to a previously mentioned item. Each text thus described a route incorporating ten items and the spatial relationships between them. The scrambled texts were prepared by presenting the title and first sentence of the story followed by the remaining nine sentences in one of four random orders. The four random orders were obtained using a random number generator and a checking procedure. The checking procedure was used to eliminate two orders correlating highly with the original order.

The sentence verification task used four types of sentences; verbatim sentences (taken from the original texts), paraphrase sentences (sentences with the same meaning as sentences in the text, but using slightly different wording), inference sentences (statements describing correct spatial relationships not explicitly described in the original texts) and distracter items (sentences based on the above items with one or more words altered to render them false).

## Procedure

The procedure and instructions were similar to that of Experiment 1. The experiment was divided into four stages. The sequence and content of the stages, along with an example of the order of presentation for one condition, is shown in Figure 3.9 below. In the first stage of the experiment each subject read one story in its unscrambled form and the other story in its scrambled form. In the second stage each subject was given an intervening sentence verification task for each story they had read. The sentences were presented, in a random order, using a personal computer. Subjects were asked to decide as quickly and accurately as possible whether each sentence was ‘true’ or ‘false’ with respect to the stories they had read by pressing one of two keys on the computer keyboard. In the third stage each subject was given both texts to reread. The instructions were to read each text through once from start to finish. Reading times for each text were recorded in both stage 1 and stage 3. In the final stage each subject was asked to draw all the landmarks and features from the places described in each of the stories they had read. They were requested to label features as they drew them. The resulting drawing orders were then recorded (as for Experiment 1). Across stages the order of the two stories (the shop text followed by the apartment text, or the apartment text followed by shop text) remains constant (refer to Figure 3.9).

Stage	Procedure carried out in this stage:
1	a) Read first text (e.g. unscrambled apartment text). b) Read second text (e.g. scrambled shop text).
2	a) Sentence verification task for first text (e.g. unscrambled apartment text). b) Sentence verification task for second text (e.g. scrambled shop text).
3	a) Reread first text (e.g. unscrambled apartment text). b) Reread second text (e.g. scrambled shop text).
4	a) Drawing recall task for first text (e.g. unscrambled apartment text). b) Drawing recall task for second text (e.g. scrambled shop text).

**Figure 3.9: The four stages of the procedure carried out in Experiment 2.**

So between reading any given story and carrying out the sentence verification task or the drawing recall task there was always an intervening activity (involving the other story). Half the subjects read or carried out a task involving an *unscrambled* story in the sub-stage labelled ‘a’. The remaining subjects read or carry out a task involving an *scrambled* story in sub-stage ‘a’. Similarly, half the subjects read the shop text in sub-stage ‘a’, while the other half read the apartment text in sub-stage ‘a’. This means that the order of presentation is counterbalanced for the type of text (unscrambled or scrambled) and the particular text used (route story 1 or route story 2).

### 3.4.2 Results

#### Reading times

There were no significant differences in the overall reading times for the different texts or between the unscrambled and scrambled text versions (though the scrambled texts did take slightly longer to read).

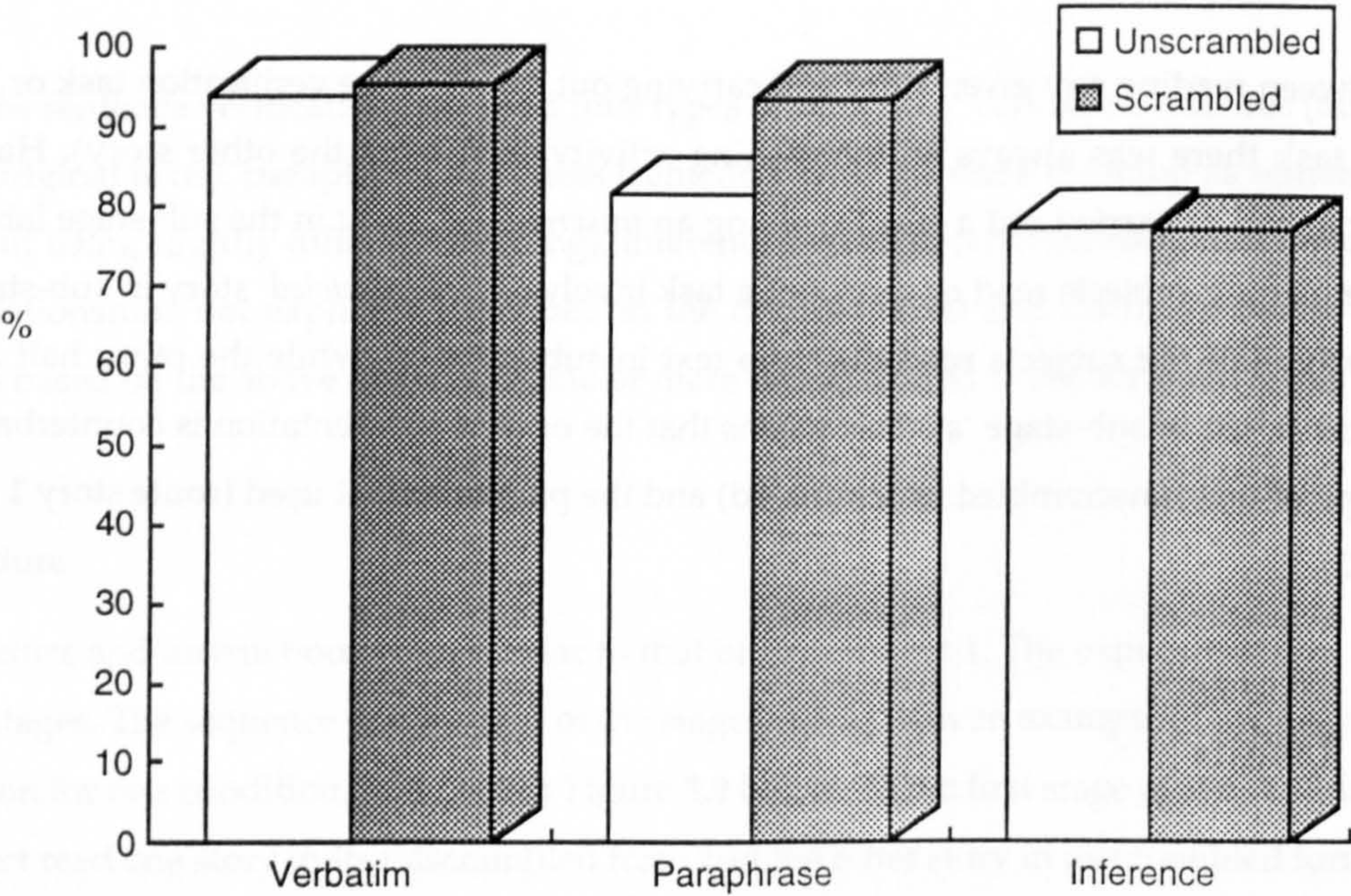
#### Performance on the sentence verification task

The overall pattern of performance on the sentence verification task between the two different stories was similar. This was confirmed using a series of Mann-Whitney U tests. There were no significant differences in total accuracy between unscrambled versions ( $U = 20.5, Z = 1.22, N = 16$ , two-tailed  $p = 0.22$ ) or the scrambled versions of the shop and apartment texts ( $U = 24, N = 16, Z = 0.85$ , two-tailed  $p = 0.40$ ). Nor were there any significant differences in response times between



the unscrambled or scrambled versions of the two stories. This was confirmed by comparing times for correct responses to verbatim, paraphrase and inference sentences. For the unscrambled text versions these comparisons were not significant for verbatim ( $U = 20$ ,  $Z = 1.26$ ,  $N = 16$ , two-tailed  $p = 0.21$ ), paraphrase ( $U = 25$ ,  $Z = 0.74$ ,  $N = 16$ , two-tailed  $p = 0.46$ ) or inference sentences ( $U = 16$ ,  $Z = 1.68$ ,  $N = 16$ , two-tailed  $p = 0.093$ ). Nor were there any significant differences for the scrambled text versions on verbatim ( $U = 32$ ,  $Z = 0$ ,  $N = 16$ , two-tailed  $p = \text{n.s.}$ ), paraphrase ( $U = 27$ ,  $Z = 0.59$ ,  $N = 16$ , two-tailed  $p = 0.56$ ) or inference sentences ( $U = 23$ ,  $Z = 0.97$ ,  $N = 16$ , two-tailed  $p = 0.33$ ). For the remaining analyses data obtained from the two stories were pooled together.

As predicted, verbatim sentences were verified more quickly and accurately than paraphrase sentences. Inference sentences were verified more slowly and even less accurately. Figures 3.10 and 3.11 show mean accuracy and response times for each type of sentence for the unscrambled and scrambled text conditions. The response times reported are only those for correct responses (though including incorrect responses does not alter the observed pattern of results).



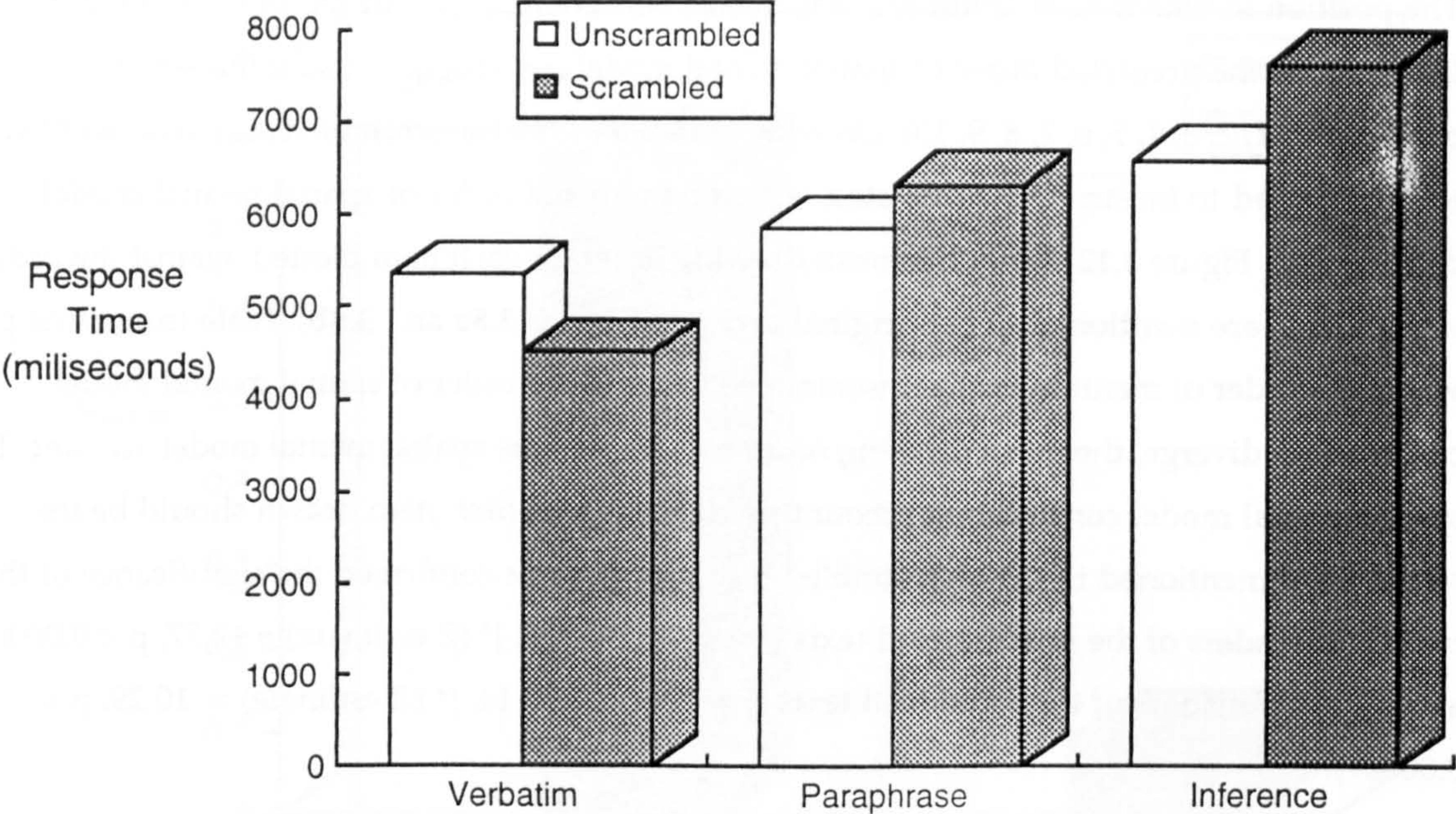
**Figure 3.10: Mean percentage accuracy on the sentence verification task for readers of unscrambled and scrambled stories**

For the accuracy data the predicted pattern of results (verbatim > paraphrase > inference) was analysed using the Page test for ordered alternatives. This analysis confirmed the significance of the observed trend for both the unscrambled ( $L = 204.5$ ,  $N = 16$ ,  $L_{crit} = 202$ ,  $p < 0.05$ ) and scrambled conditions ( $L = 207$ ,  $N = 16$ ,  $L_{crit} = 202$ ,  $p < 0.05$ ).

For the response times, the predicted pattern of results (inference > paraphrase > verbatim) was analysed in the same way. This analysis confirmed the significance of the observed trend for both the unscrambled ( $L = 205$ ,  $N = 16$ ,  $L_{crit} = 202$ ,  $p < 0.05$ ) and scrambled conditions ( $L = 209$ ,  $N = 16$ ,  $L_{crit} = 202$ ,  $p < 0.05$ ). The median difference in response times between



inference and paraphrase sentences is relatively small (the mean differences are somewhat inflated by outliers).



**Figure 3.11: Mean response times on the sentence verification task for readers of unscrambled and scrambled stories (correct responses only)**

However, it should be noted that the inference statements are, on average, shorter in length (and hence reading time) than the paraphrase statements. The paraphrase sentences were created from the verbatim sentences and so were matched for length. There is an indication in the two graphs that the scrambled text conditions produced better and faster recall for verbatim information and better recall for paraphrases. This trend is does not reach statistical significance.

**Levels of recall on the drawing task**

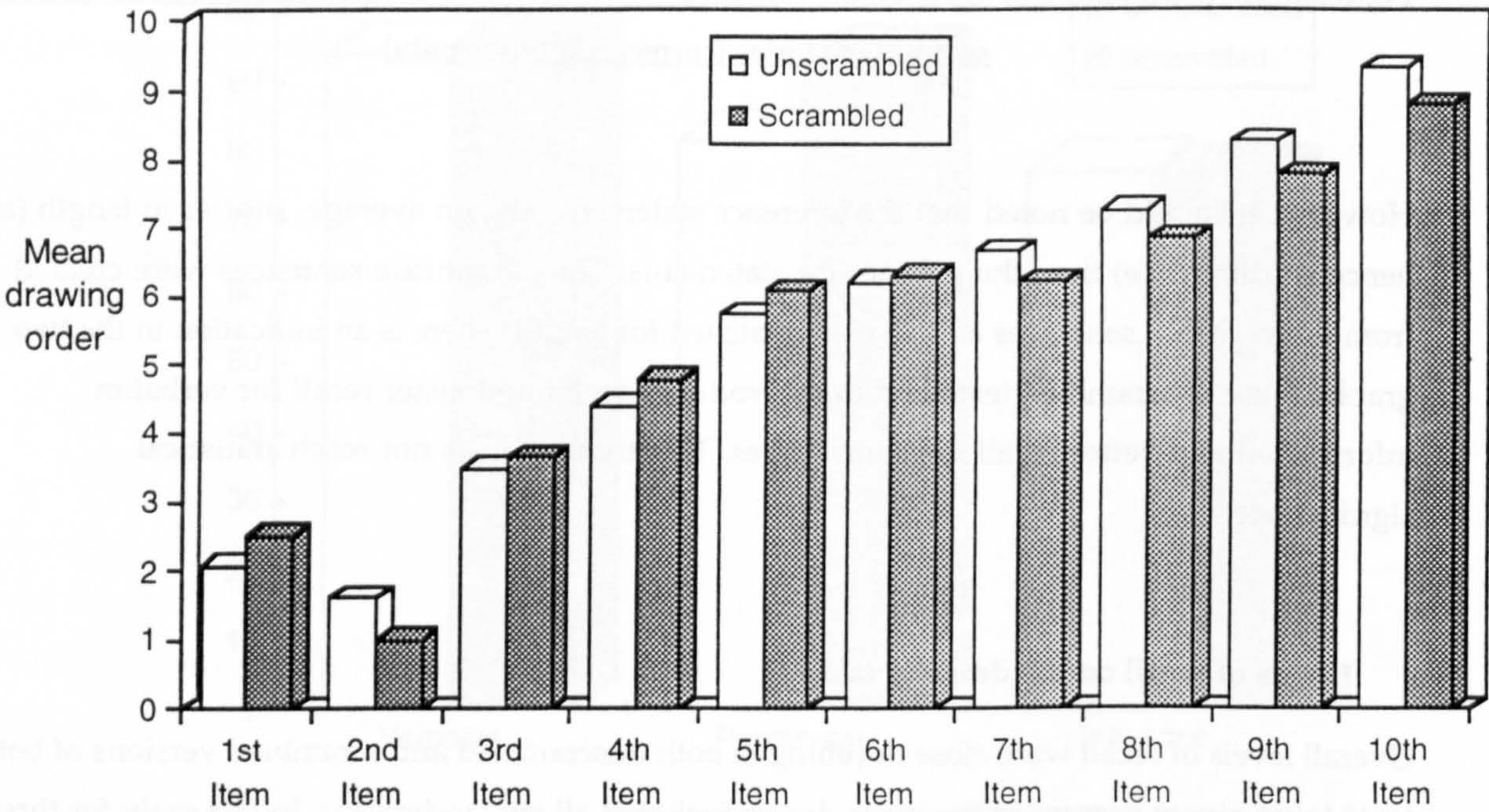
Overall levels of recall were close to ceiling in both unscrambled and scrambled versions of both texts (over ninety percent of the maps drawn included all ten landmarks). In fact, only for three drawings (one for an unscrambled text and two for a scrambled text) were any items omitted. These three drawings were therefore not included in the subsequent analysis (though the drawing order of the remaining landmarks also tended to confirm to the drawing orders predicted in the introduction).

**Drawing order**

There were no significant differences in recall or drawing order between the two different locations described in the texts. Therefore for the remaining analyses data from the two different original stories were pooled together (in fact, the significant differences reported below are also found if the two stories are analyzed separately). Drawing order was recorded for



each of the items in the analysis for readers of both the scrambled and unscrambled narratives. The data was analyzed between items using the Jonckheere trend test (Siegel & Castellan, 1988). The position in which each landmark was drawn was compared with the position of that landmark in the predicted order of spatial mental model construction (using the set of coefficients: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10). Drawing orders for both the scrambled and unscrambled texts appeared to be strongly associated with the predicted order of spatial mental model construction. Figure 3.12 shows the mean drawing order for each item plotted against the order in which items are mentioned in the original text (see Figures 3.8a and 3.8b). Note that at the point where the order of mention in the unscrambled text and the order of spatial mental model construction diverge, the mean drawing order conforms to the spatial mental model account. The spatial mental model construction account predicts that the *first* item drawn should be the *second* item mentioned in the unscrambled text. The analysis confirmed the significance of these trends for readers of the unscrambled texts ( $J = 8971, N = 15, J^* (Z \text{ estimate}) = 12.77, p < 0.00001$ ) as well as for readers of the scrambled texts ( $J = 7251.5, N = 14, J^* (Z \text{ estimate}) = 10.29, p < 0.00001$ ).

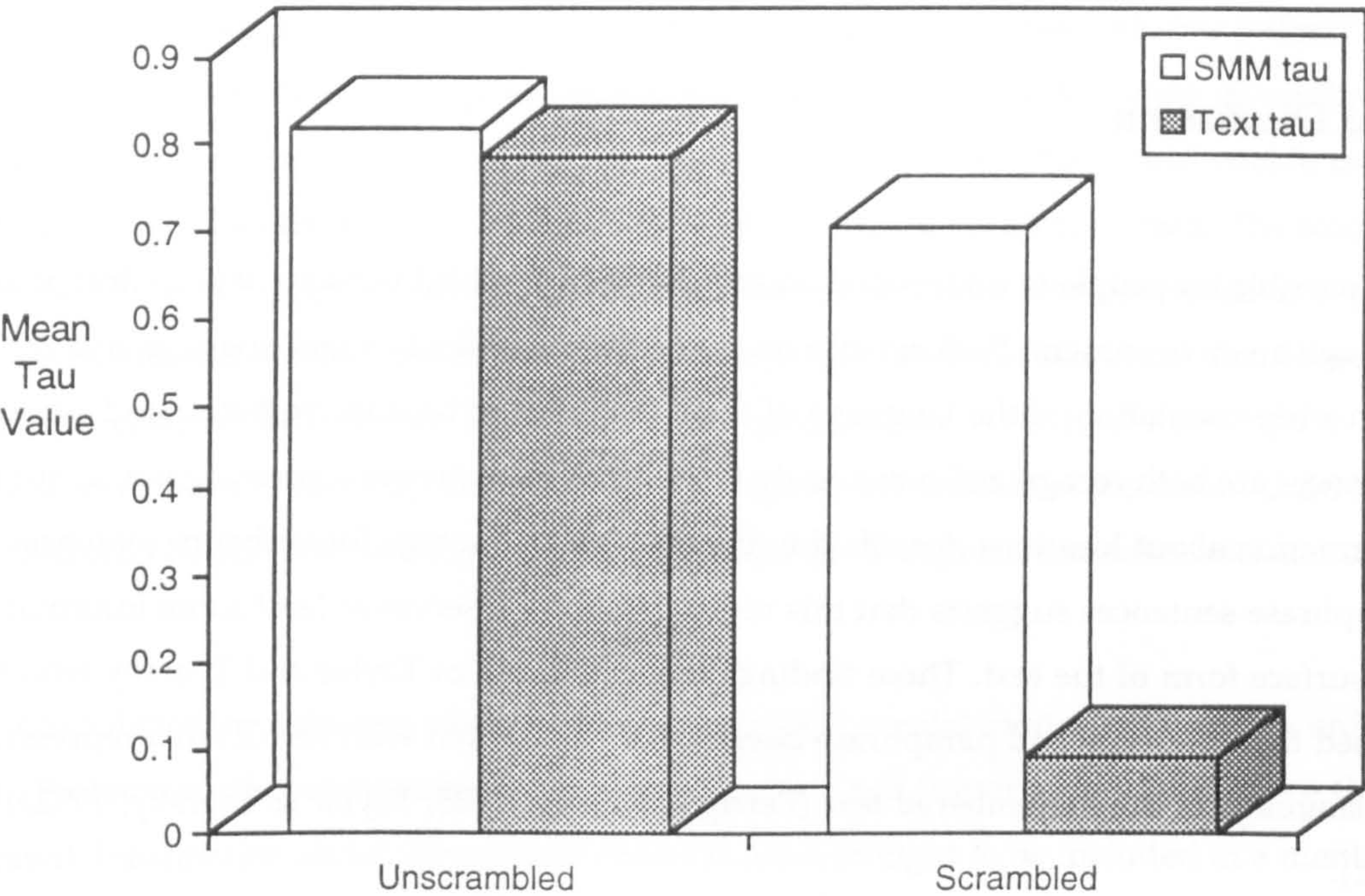


**Figure 3.12: Mean drawing order against order of occurrence in the unscrambled text for readers of the scrambled and unscrambled texts.**

Kendall's tau was calculated for each drawing order with the predicted spatial mental model construction order as well as for the actual order of occurrence of each item in the text that subjects had read. Figure 3.13 shows the mean tau values with spatial mental model construction order (SMM tau) and text order (text tau) for the readers of both scrambled and unscrambled stories. For readers of the scrambled texts tau values with spatial mental model construction order were significantly higher than those for text order (Wilcoxon's  $T = 2, N = 14, Z = 3.17, p =$

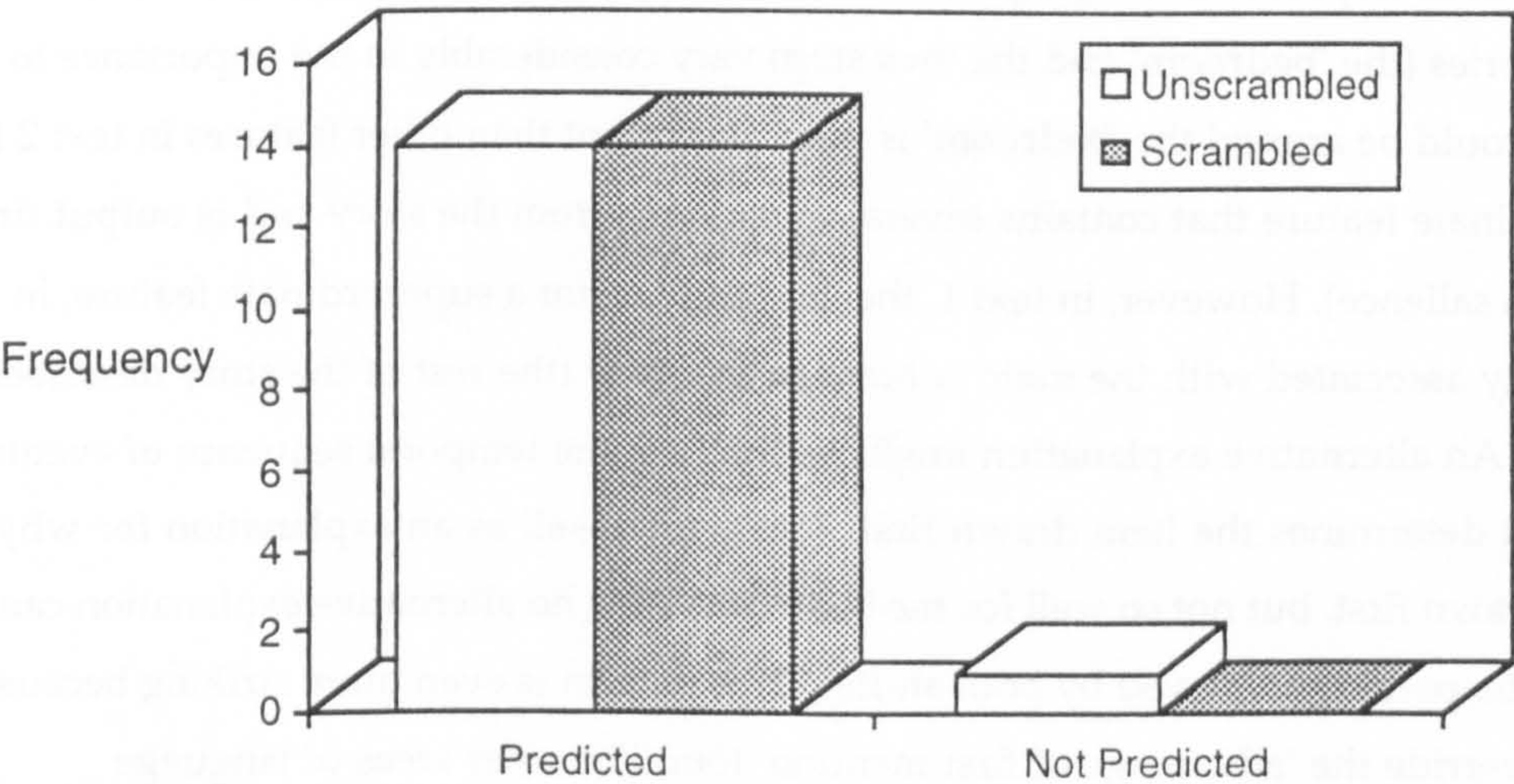


0.015). However, for the unscrambled text there was no significant difference in tau values with text order and the tau values with spatial mental model construction order.



**Figure 3.13: Mean concordance with text and spatial mental model construction order for readers of both unscrambled and scrambled texts**

The final analysis of drawing order looked only at the first item drawn by subjects. The spatial mental model account predicts that the item mentioned second in each text should be drawn first (depending on the text this should be either the 'bedroom' or the 'bus stop'). Figure 3.14 shows the relative frequency with which the predicted item is drawn first:



**Figure 3.14: Relative frequency with which the item predicted by the spatial mental model account is drawn first for the unscrambled and scrambled text versions.**

This pattern of results was analyzed using the binomial test (Siegel & Castellan, 1988). For the unscrambled stories fourteen subjects drew the 'second' item first ( $N = 15$ ,  $p < 0.0001$  using the



binomial test). For the scrambled stories all fourteen readers drew the 'second' item first ( $N = 14$ ,  $p < 0.0001$ ).

### 3.4.3 Discussion

It is possible for people to understand and remember scrambled versions of texts describing routes through an environment. Performance on the sentence verification task suggests that readers do form a representation of the language of the text they read because verbatim and paraphrase sentences are both recognized more easily than inference sentences containing novel, but true, information about locations described in the text. The advantage for verbatim sentences over paraphrase sentences suggests that this representation preserves at least some information about the surface form of the text. These findings replicate those of Taylor and Tversky who have argued that inference and paraphrase questions are answered with respect to a representation of the language of the remembered text (Perrig & Kintsch, 1985; Taylor & Tversky, 1992a; Tversky, 1991).

Drawing recall orders for scrambled and unscrambled spatial descriptions are significantly related to the predicted order of spatial mental model construction. In particular, the first item drawn appears to be very strongly influenced by the predicted order of spatial mental model construction rather than the order items are mentioned in the description. Because four different scrambled text orders were used it is unlikely that these results for the scrambled texts could be due to a particular text order. It is possible that the item drawn first is an artefact of a factor such as importance. This explanation is unlikely because the items mentioned second in the two stories (the 'bedroom' and the 'bus stop') vary considerably in the importance to the narrative. It could be argued the 'bedroom' is more important than other features in text 2 (e.g. it is a superordinate feature that contains several other items from the story and is output first because of its salience). However, in text 1, the 'bus stop' is not a superordinate feature, in fact, it is only weakly associated with the main content of the story (the rest of the story describes a music shop). An alternative explanation might be that it is the temporal sequence of events from the story that determines the item drawn first. This works well as an explanation for why the bus stop is drawn first, but not so well for the bedroom. Thus no alternative explanation can account for the pattern displayed by both stories. This pattern is even more striking because it appears to override the 'advantage of first mention' found in other areas of language comprehension (Gernsbacher, 1990; 1991).

Two possible criticisms of this experiment also need to be addressed. The first criticism is that the sentence verification task carried out by subjects may have contaminated recall on the drawing task. This is a weak criticism for a number of reasons. The order of the trials in the

sentence verification task was randomly determined by the computer for each subject. It is therefore virtually impossible that the order in which sentences were presented could have introduced any systematic bias. In fact, any contamination should increase between subject error and reduce the power of any subsequent statistical test. In addition, subjects were given the opportunity to reread each story after the sentence verification task. This would enable them to take account of any incorrect responses made during the sentence verification task. The second criticism is that the findings presented here rest too heavily on only a small deviation between the predicted order of spatial mental model construction and the order items are mentioned in the text. This criticism is true only in the case of the unscrambled texts. In the case of the scrambled texts there is virtually no relationship between text order and the actual drawing orders produced by subjects. In both cases the results favour the spatial mental model account.

On balance the evidence seems to favour the spatial mental model construction order account. Perhaps spatial relationships, temporal sequence and importance all influence the first item drawn. Information about the order of events is often thought to be included in a mental model and in particular within a spatial mental model (Bower & Morrow, 1990; Glenberg & Langston, 1992; Johnson-Laird, 1983). On the other hand it is very difficult to see how a representation of the language of the text could account for these influences. A mental model is a representation of the situation described by a text and hence should be influenced by the spatial and temporal properties of the situation as well as by general world knowledge brought to bear by the reader (Johnson-Laird, 1983; van Dijk & Kintsch, 1983).



### 3.5 Experiment 3

Experiments 1 and 2 have replicated the drawing order effect obtained by Taylor and Tversky (1992a). Experiment 2 demonstrated that the same effect is observed when readers of scrambled narratives are asked to draw maps of the places they have read about. It has been argued that these findings are best explained by memory for a spatial mental model of the situation described by the text rather than by memory for the language of the text. One possibility, however, has not fully been discounted. In both the previous experiments readers were presented with the entire text and were able to refer back to previous sections of the text as they read. It is possible, if unlikely, that readers were able to identify the correct order of the scrambled narratives with only *minimal* reading of the scrambled text itself. Identifying at least some of the original sequence of the scrambled text is an essential part of spatial mental model construction (this point is discussed in more detail at the end of this chapter in section 3.4.2). However, people might be able to 'reorder' the scrambled text without ever reading the text in its scrambled order, provided they are able to capitalize on superficial cues (e.g. identifying the first sentence and then scanning through the text for subsequent sentences which mention the same item or items). This possibility is eliminated in Experiment 3 by introducing a new method of presenting readers with the unscrambled and scrambled texts. Each text is presented one sentence at a time and readers are prevented from referring back to previous sections of the text. This manipulation prevents readers from using 'superficial' cues to reorder the scrambled narratives.

Two other major differences were introduced into the experimental design. As in Experiment 2, two short route narratives were used. The first story, a short fairy tale used only asymmetric spatial relations (in this case North, South, East and West). The second story used a mixture of asymmetric (e.g. left and right) and symmetric spatial relations (e.g. opposite or across from). As noted in Chapter 2 the spatial mental model account makes different predictions about the first item entered into a spatial mental model according to whether the first spatial relationship described is symmetrical or asymmetrical. It is predicted that where the first spatial relationship is asymmetrical (as in Experiment 2) the second item mentioned in the description is entered as the initial token in the spatial mental model and acts a reference point for the first item. Where the first spatial relationship in a description is symmetric neither of the two items acts as a reference point and readers are likely to accept the first item mentioned as the initial token in their spatial mental model. Figure 3.15 shows the first sentences of both the stories used in Experiment 3 (see Figures 3.16a and 3.16b for the complete stories). The second change to the experimental design was that each reader was not only asked to draw the place described in each of the two stories, but also to write out as much as they could remember of the six sentences comprising the story. Half the readers drew the story before writing it out and half the readers wrote the story before being asked to draw it.

Story A (first sentence uses an asymmetric spatial relation)

Princess Jane lived in an *Ancient Castle* on the Southern shore of an *Enchanted Lake*.

Story B (first sentence uses a symmetric spatial relation)

Mr. Simpson lives in a *Thatched Cottage* immediately opposite the old, red *Telephone Box*.

Figure 3.15: The first sentences of the two stories presented to subjects in Experiment 3.

The spatial mental model account predicts that the drawing orders of all the readers will preserve the order of the original, unscrambled story. One exception to this is predicted. Readers of Story A should draw the second mentioned item first because of the asymmetry of the relation “on the Southern shore of” (see Figure 3.15). Readers of story B should draw all the landmarks in the order they are mentioned in the unscrambled narrative. For the story recall it is predicted that verbatim memory for the two stories will be poor because readers will attempt to construct a coherent spatial mental model and may not actively attempt to remember the language of the text.

### 3.5.1 Method

#### Subjects

Nineteen graduate students and members of staff from the Open University participated in the study. The age of the subjects ranged from early twenties to early forties.

#### Design

There were two within subject variables; text organization (each subject read both an *unscrambled* and a *scrambled* narrative) and the type of spatial relation employed in the first sentence of each text (*asymmetric* in story A or *symmetric* in story B). The order of the scrambled narrative (one of two random sequences of sentences) varied between subjects. Text organization and the location described by the text were counterbalanced. Subjects were assigned randomly to one of the counterbalanced conditions and to one of the two scrambled text orders. The dependent variables were whether and in what order a landmark was recalled in either of the two free recall tasks; drawing the place described by the text or writing out the sentences of the text.



Story A (unscrambled).

Princess Jane lived in an *Ancient Castle* (1) on the Southern shore of an *Enchanted Lake* (2).

One day she left the castle and travelled West along the *Great High Road* (3) in search of adventure.

After travelling for a day along the Great High Road she came to *Fire Mountain* (4).

Near the summit of Fire Mountain Princess Jane was able to see the entrance to *The Dragon's Cave* (5) where Prince Alan was being held captive.

Princess Jane rescued the young Prince from the Dragon's Cave and they escaped Northwards through a *Secret Tunnel* (6).

When they emerged from the Secret Tunnel they were able to walk East to the safety of the *Imperial Palace* (7), where they lived happily ever after.

Figure 3.16a: The unscrambled version of story A.

Story B (unscrambled).

Mr. Simpson lives in a *Thatched Cottage* (1) immediately opposite the old, red *Telephone Box* (2).

Every Sunday Mr. Simpson leaves his home and walks next door to the *Village Shop* (3).

After having bought the Sunday papers from the Village Shop he crosses over the road and comes to the *Village Common* (4).

On the far side of the Village Common lies the *Cricket Pavilion* (5) where Mr. Simpson pauses and watches the local team prepare for a match.

Mr. Simpson then walks behind the Cricket Pavilion and comes to *Church Lane* (6).

At Church Lane he turns left and walks down the road until he comes to the *White Horse Inn* (7), where he stops for lunch.

Figure 3.16b: The unscrambled version of story B.

**Materials**

The two stories used in the experiment (Story A and Story B) are presented in Figures 3.16a and 3.16b. Both unscrambled texts were six sentences long. Each sentence contained a spatial relationship between two items. Each sentence after the first introduced one new landmark. In story A the first sentence used an asymmetric spatial relation, whereas in story B the first sentence used a symmetric spatial relation (refer to Figure 3.15). The scrambled texts were prepared by presenting the six sentences in one of two random orders. The two random orders were obtained using a random number generator and a checking procedure. The checking procedure was used to eliminate any orders correlating significantly with the original order.

**Procedure**

The procedure and instructions were similar to that of Experiment 2. The experiment was divided into three stages. The procedure for each stage is laid out in Figure 3.17 below. In the first stage of the experiment each subject read one unscrambled and one scrambled story. In the second stage half the subjects were asked to draw the places described in both the original stories (the drawing recall task) while the remaining half were asked to write out the six sentences that comprised the original stories (the story recall task). In the third stage subjects who had carried out a drawing recall task in stage 2 carried out the story recall task. The remaining subjects carried out the drawing recall task in stage 3.

Stage	Procedure carried out in this stage:
1	a) Read unscrambled text A <i>or</i> scrambled text B. b) Read scrambled text B <i>or</i> unscrambled text A.
2	<i>either</i> i) Drawing recall task <i>or</i> ii) Story recall task.
3	<i>either</i> i) Story recall task <i>or</i> ii) Drawing recall task.

**Figure 3.17: The three stages of the procedure for Experiment 3.**

In stage 1 the instructions and both story A and story B were presented using Hypercard on a Macintosh personal computer. Each story (unscrambled or scrambled) was presented one sentence at a time. Each reader was allowed as much time as they wished to read each sentence, but each sentence was only presented once. If a reader had been presented with story A first he or she would attempt to recall story A first in both stage 2 and stage 3. If a reader had been presented with story B first he or she would attempt to recall story B first in both stage 2 and stage 3. Thus each subject read one unscrambled text and one scrambled text carried out both the story recall task and the drawing recall task for each text. It was not possible for subjects to rely

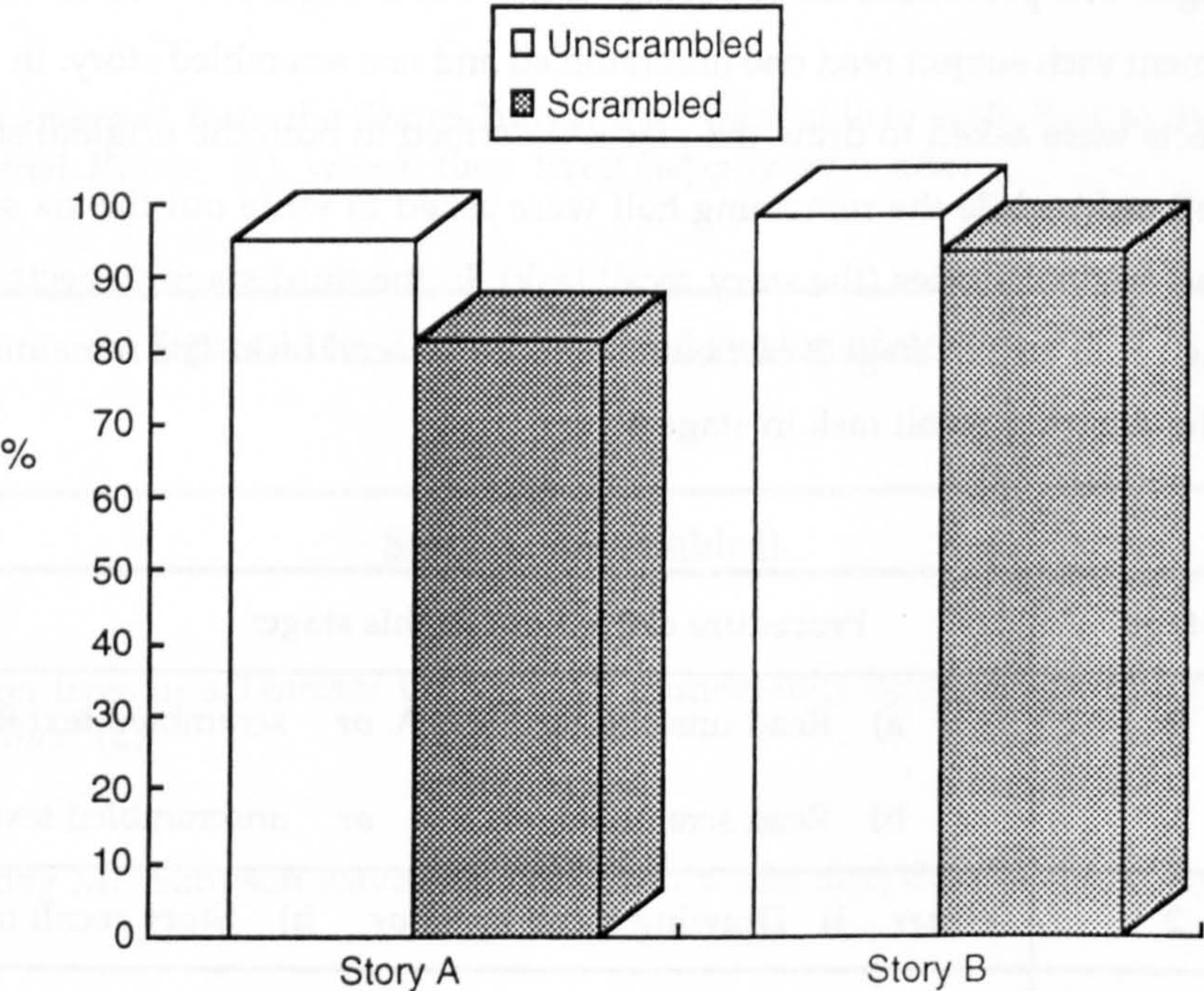


on recall from working memory to remember either story, because there was always an intervening task (reading or recalling the other story).

### 3.5.2 Results

#### Levels of recall on the drawing task

Overall levels of recall on the drawing task for the both the unscrambled and scrambled stories were almost at ceiling (see Figure 3.18). Only those readers who successfully recalled at least six of the seven landmarks of a given story were included in the analysis. Out of the total of 38 drawings only 5 could not be included in the analysis because they contained fewer than six correct landmarks. In the small number of cases where only six landmarks were recalled it was assumed that the missing item would be recalled last.

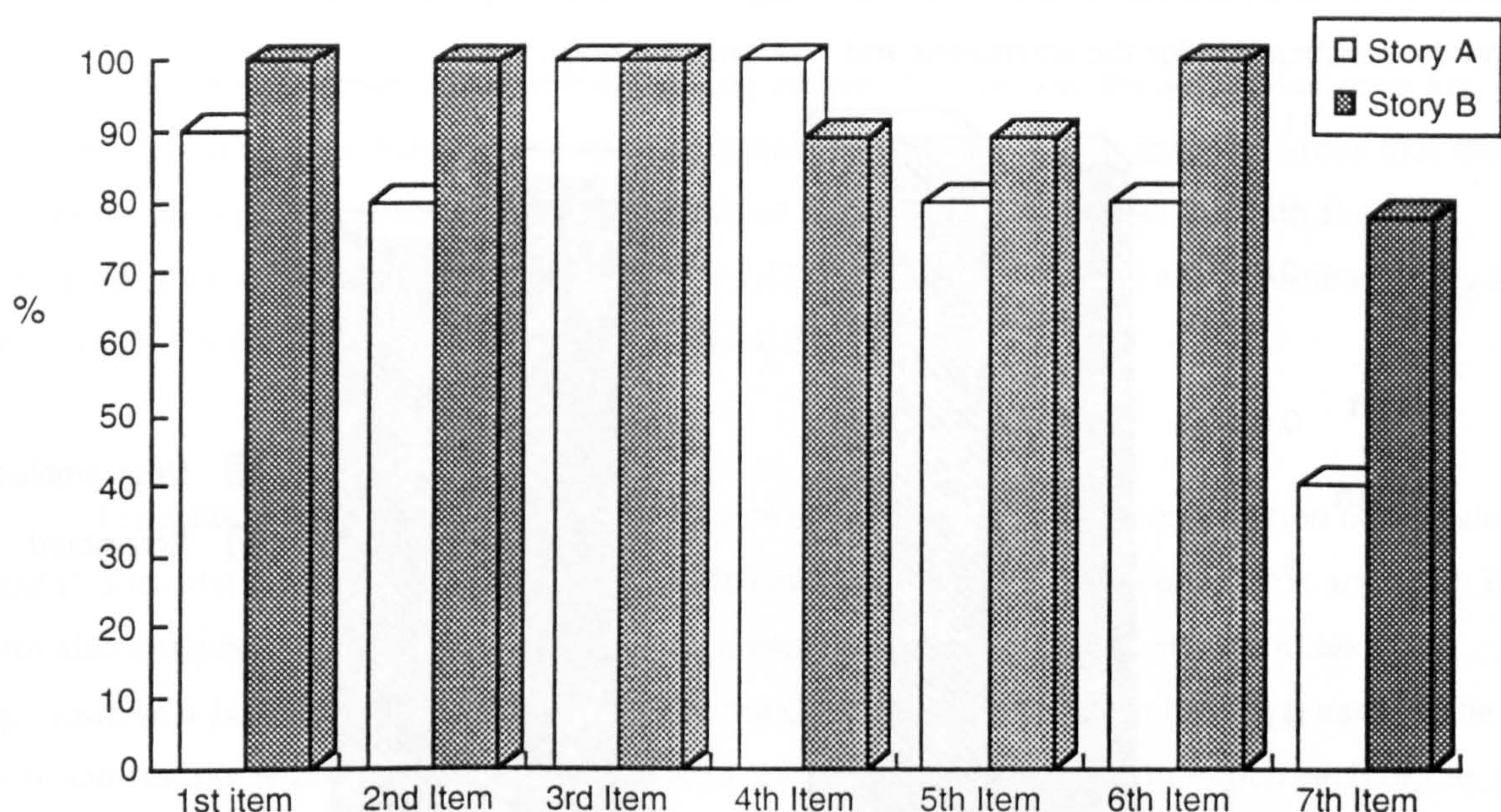


**Figure 3.18: Mean percentage levels of recall for unscrambled and scrambled texts by story.**

#### Recall of landmarks (scrambled texts only)

As has been noted drawing recall of landmarks was at ceiling for the unscrambled versions of both stories. Levels of recall for the scrambled versions of the two stories was analyzed using a non-parametric trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). Each set of data was tested (using the set of coefficients: +3, +2, +1, 0, -1, -2, -3) for a trend against the predicted order of spatial mental model construction and separately against the original order landmarks were mentioned in the appropriate scrambled text. Figure 3.19 shows the pattern of recall for the scrambled versions of both story A and story B (the landmarks are numbered 1 to 7 in the predicted order of spatial mental model construction):





**Figure 3.19: Mean percentage recall of landmarks for story A and story B against predicted order of spatial mental model construction (scrambled texts only).**

The trend tests for both story A ( $N = 10$ ,  $Z$  estimate = 2.64, one-tailed  $p = 0.0041$ ) and story B ( $N = 9$ ,  $Z$  estimate = 1.83, one-tailed  $p = 0.034$ ) were significant. These results provide some support for the prediction that recall is influenced by the order of spatial mental model construction. In neither of the two stories does recall for the second item entered into the spatial mental model exceed that for the first (though in both cases recall for the first two items is close to ceiling making a direct comparison difficult). That both trend tests reach significance is probably due to the large difference in recall (50% and 22% respectively) between the first and last items (the two items weighted most heavily by the analysis). For this reason the two tests should be interpreted with caution.

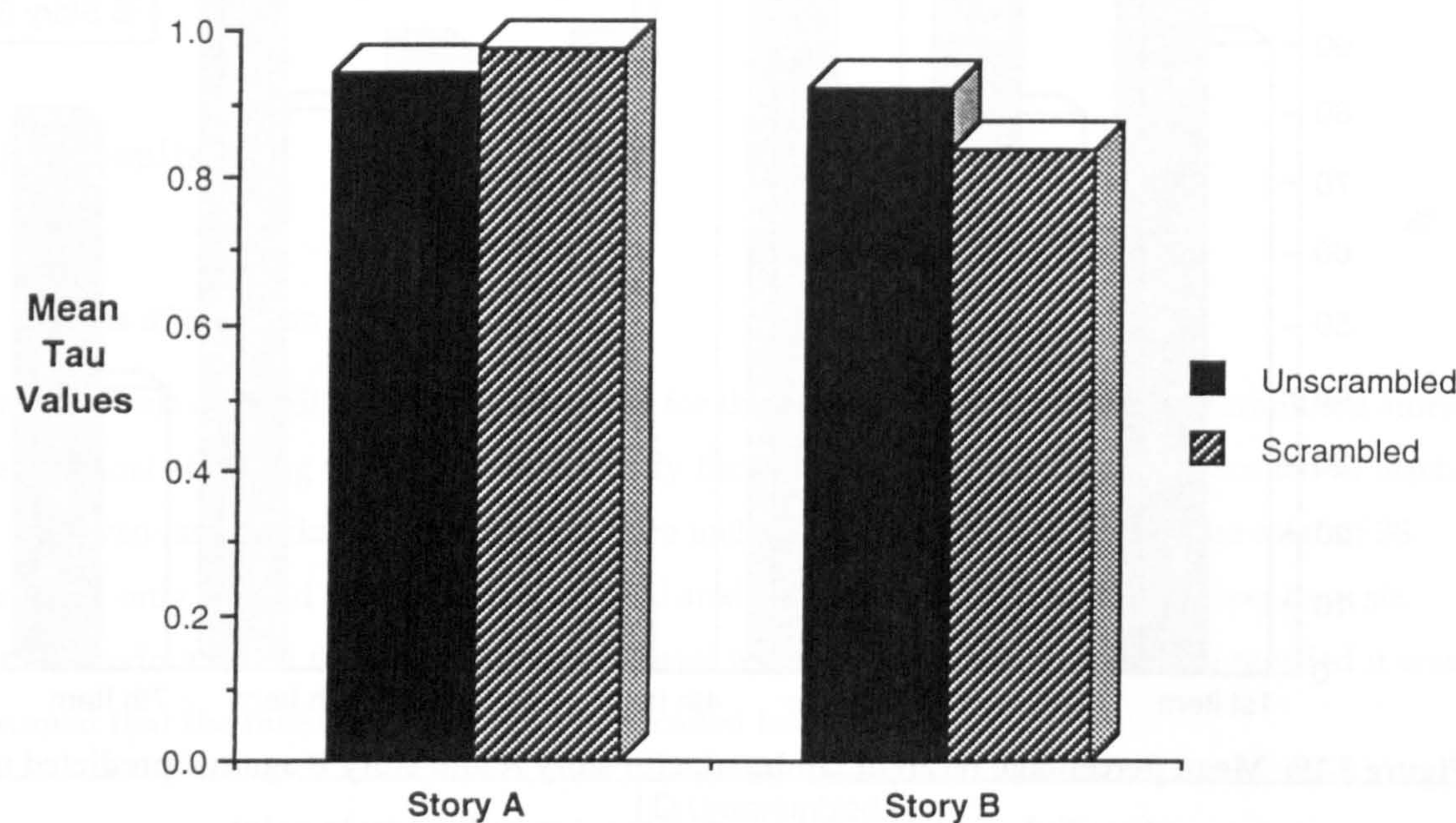
The results of the trend tests performed against the scrambled order in which the landmarks were presented were less clear cut. For story A there was no significant relationship between the 'scrambled' order and landmark recall ( $N = 10$ ,  $Z$  estimate = 0.26, one-tailed  $p = 0.40$ ). For story B there was a significant negative relationship between the 'scrambled' order and the level of recall for the seven landmarks ( $N = 9$ ,  $Z$  estimate = -1.91, one-tailed  $p = 0.028$ ). This may be an artefact of a small negative relationship between spatial mental model construction order and the order of some of the scrambled text. There is no evidence that items mentioned early in either of the scrambled texts are remembered better than items mentioned later on.

### Drawing order

As in previous experiments Kendall's tau was used as a measure of the relationship between subjects drawing orders and the predicted order of spatial mental model construction. Figure 3.20

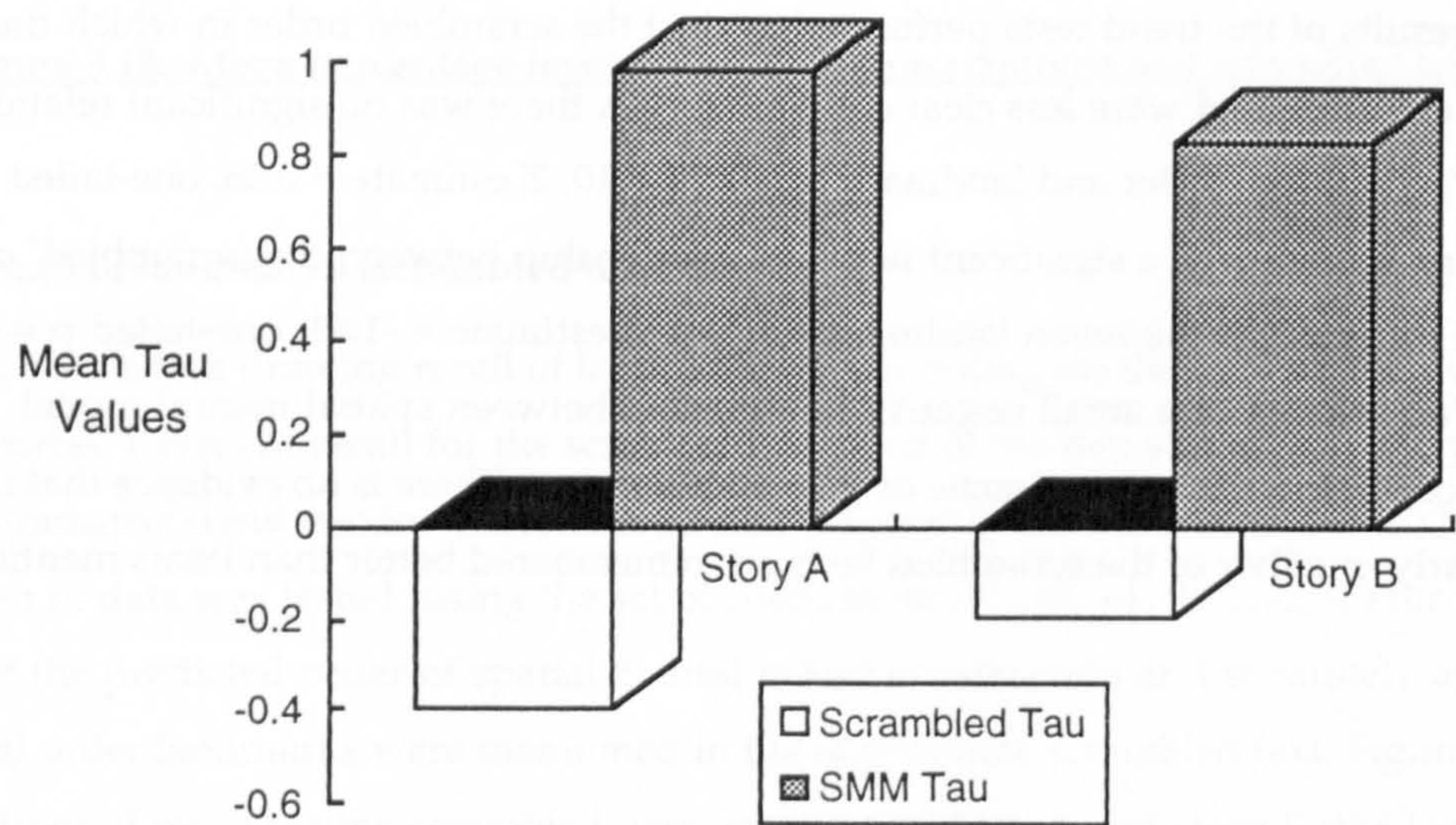


shows the mean tau values for subject drawing orders and the predicted order of spatial mental model construction for the scrambled and unscrambled texts.



**Figure 3.20: Mean tau values with predicted order of spatial mental model construction for unscrambled and scrambled texts by story.**

Tau values for story A and story B are shown separately as the predicted order of spatial mental model construction differs between the two (2, 1, 3, 4, 5, 6, 7 for story A and 1, 2, 3, 4, 5, 6, 7 for story B). None of the differences in tau values shown in Figure 3.20 reached significance. For the scrambled texts drawing orders were also compared with the order that landmarks were mentioned when the text was presented in the first stage of the experiment. Figure 3.21 shows Kendall's tau between the drawing orders of the subjects and the actual order that the landmarks were presented in (scrambled tau). Tau values with the predicted order of spatial mental model construction (SMM tau) are given for comparison.

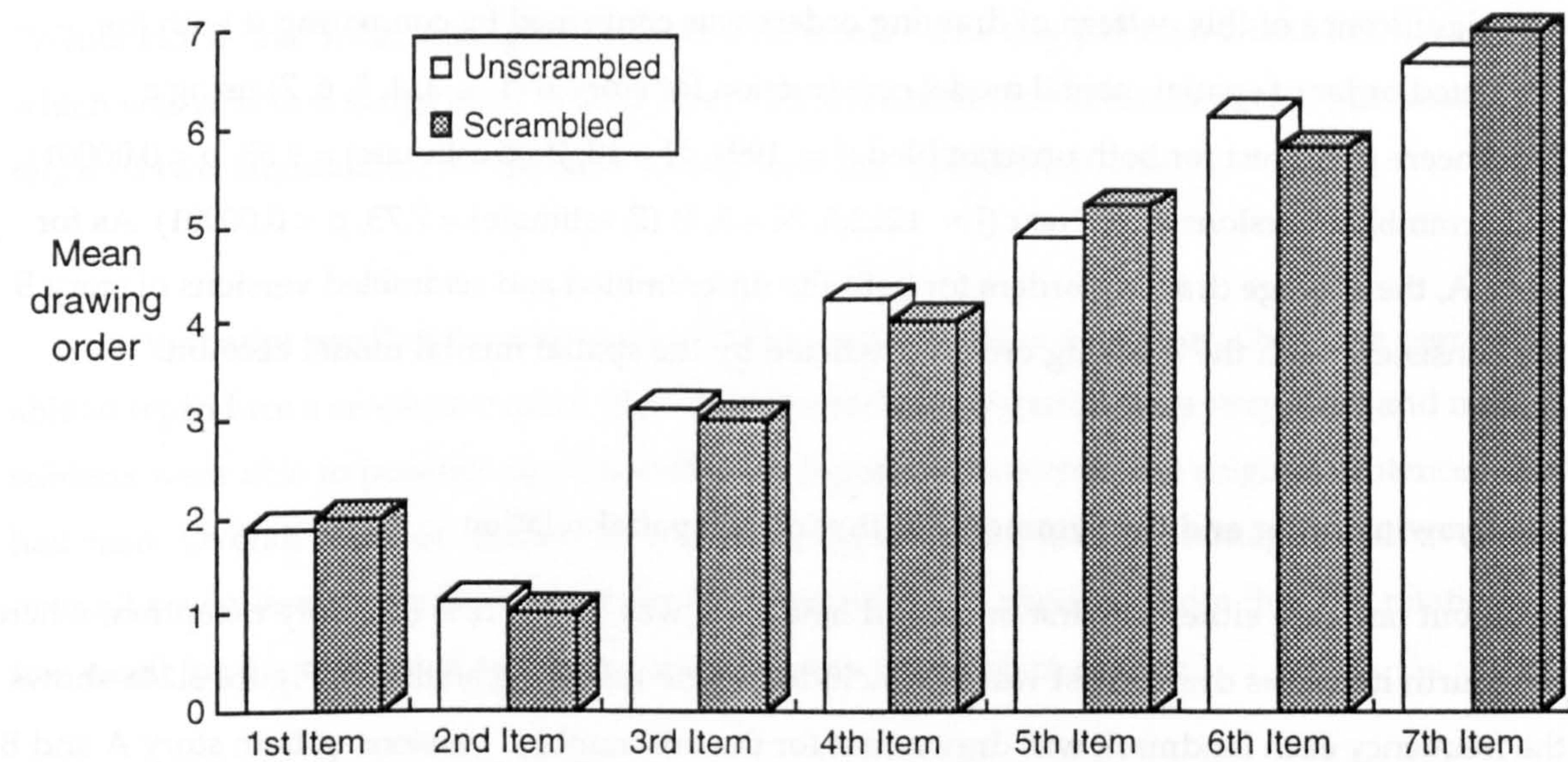


**Figure 3.21: Mean tau values for scrambled texts by story with the scrambled order of presentation and the order of spatial mental model construction.**



Analysis of these data confirms that the drawing orders of readers of the scrambled texts are more strongly related to the order of spatial mental model construction than the order that the landmarks are actually presented in the scrambled text. This was the case for both the scrambled story A (Wilcoxon's  $T = 0$ ,  $N = 5$ ,  $Z = 2.12$ , two-tailed  $p = 0.034$ ) and scrambled story B (Wilcoxon's  $T = 0$ ,  $N = 8$ ,  $Z = 2.55$ , two-tailed  $p = 0.011$ ).

In addition to comparing concordance with spatial mental model construction order using Kendall's tau the drawing orders for unscrambled and scrambled versions of story A and story B were also analyzed by item. For each of the four possible combinations of story and text organization a Jonckheere trend test was performed on the serial position in which each of the seven landmarks was drawn. Figure 3.22 shows the mean serial position in which each of the seven landmarks (numbered 1 to 7 in order of presentation in the unscrambled story version) of story A was drawn for both the unscrambled and scrambled text version.

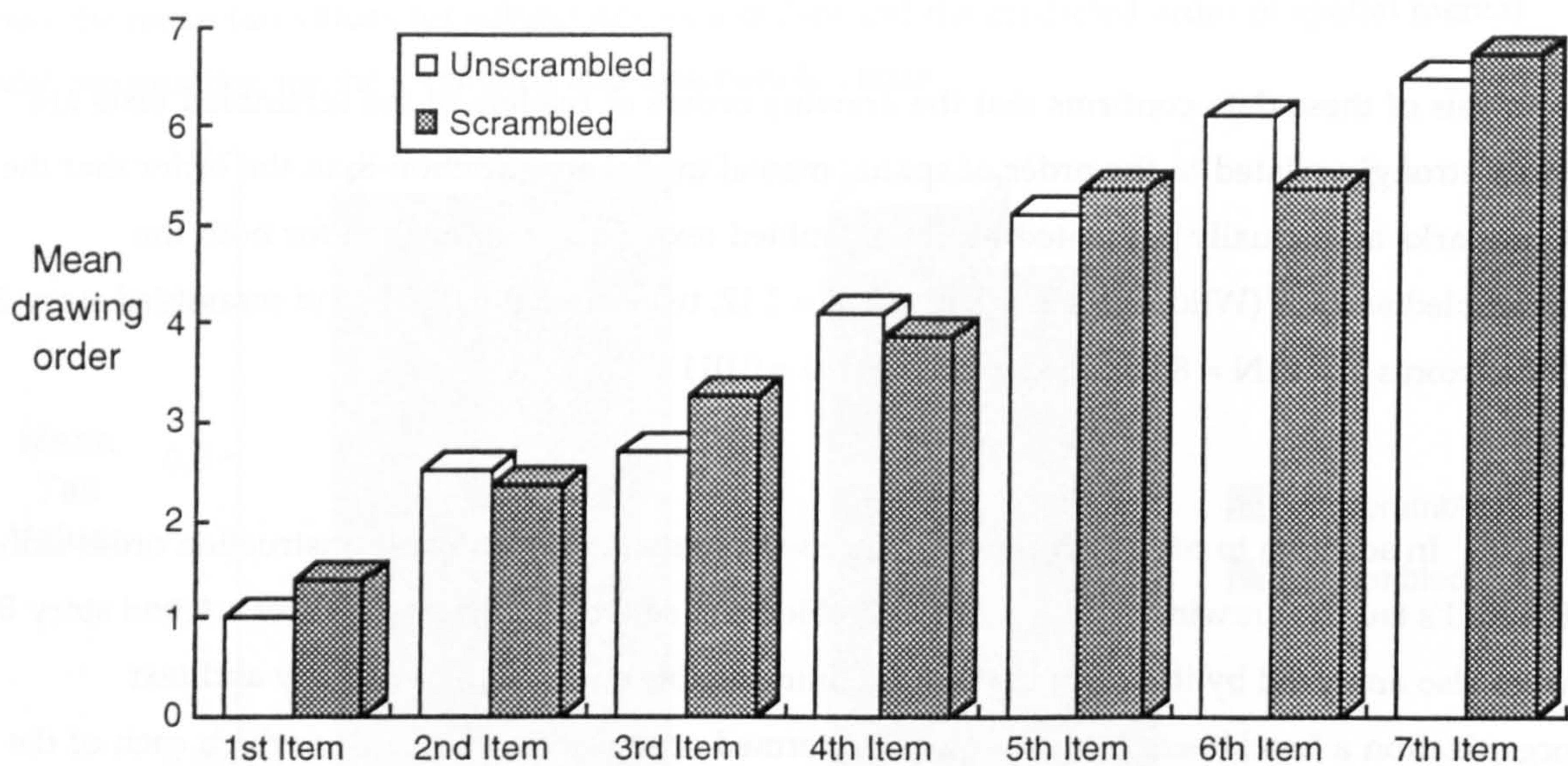


**Figure 3.22: Mean drawing order for each of the seven items from Story A.**

The significance of this pattern of drawing orders was confirmed by comparing it between items with the predicted order of spatial mental model construction for story A (2, 1, 3, 4, 5, 6, 7) using a Jonckheere trend test for both unscrambled ( $J = 1636.5$ ,  $N = 9$ ,  $J^*$  (Z estimate) = 9.43,  $p < 0.00001$ ) and scrambled versions of the text ( $J = 520$ ,  $N = 5$ ,  $J^*$  (Z estimate) = 7.40,  $p < 0.00001$ ). The mean drawing order for each of the seven landmarks in story A conforms exactly to the predicted order of spatial mental model constructions. Even when the story has been scrambled subjects tend to draw the landmarks in the order predicted by the spatial mental model account (the spatial mental model construction order for the *unscrambled* text).

The same analysis was performed on the drawing orders obtained from readers of story B. Figure 3.23 shows the mean drawing orders for each of the seven landmarks for story B:





**Figure 3.23: Mean drawing order for each of seven items from Story B.**

The significance of this pattern of drawing orders was confirmed by comparing it with the predicted order of spatial mental model construction for story B (1, 2, 3, 4, 5, 6, 7) using a Jonckheere trend test for both unscrambled ( $J = 1994$ ,  $N = 10$ ,  $J^*$  (Z estimate) = 9.68,  $p < 0.00001$ ) and scrambled versions of the text ( $J = 1212.5$ ,  $N = 8$ ,  $J^*$  (Z estimate) = 7.73,  $p < 0.00001$ ). As for story A, the average drawing orders for both the unscrambled and scrambled versions of story B are consistent with the drawing orders predicted by the spatial mental model account.

### Drawing order and the symmetry of the initial spatial relation

In all but one case either the first or second landmark was drawn first (the only exception, where the fourth item was drawn first was not included in the following analysis). Figure 3.24a shows the frequency each landmark was drawn first for the unscrambled versions of both story A and B. Figure 3.24b shows the frequency each landmark is drawn first for the scrambled versions of story A and B.

	Unscrambled texts	
	Landmark 1	Landmark 2
Story A	1	8
Story B	9	0

**Figure 3.24a**

	Scrambled texts	
	Landmark 1	Landmark 2
Story A	0	5
Story B	7	0

**Figure 3.24b**

These frequency data appear to support the predicted pattern of results. This is the prediction that an asymmetric initial spatial relation (as used in story A) will result in the landmark mentioned second being drawn first while a symmetric initial spatial relation (as in story B) will result in the first landmark mentioned being drawn first. The significance of this pattern

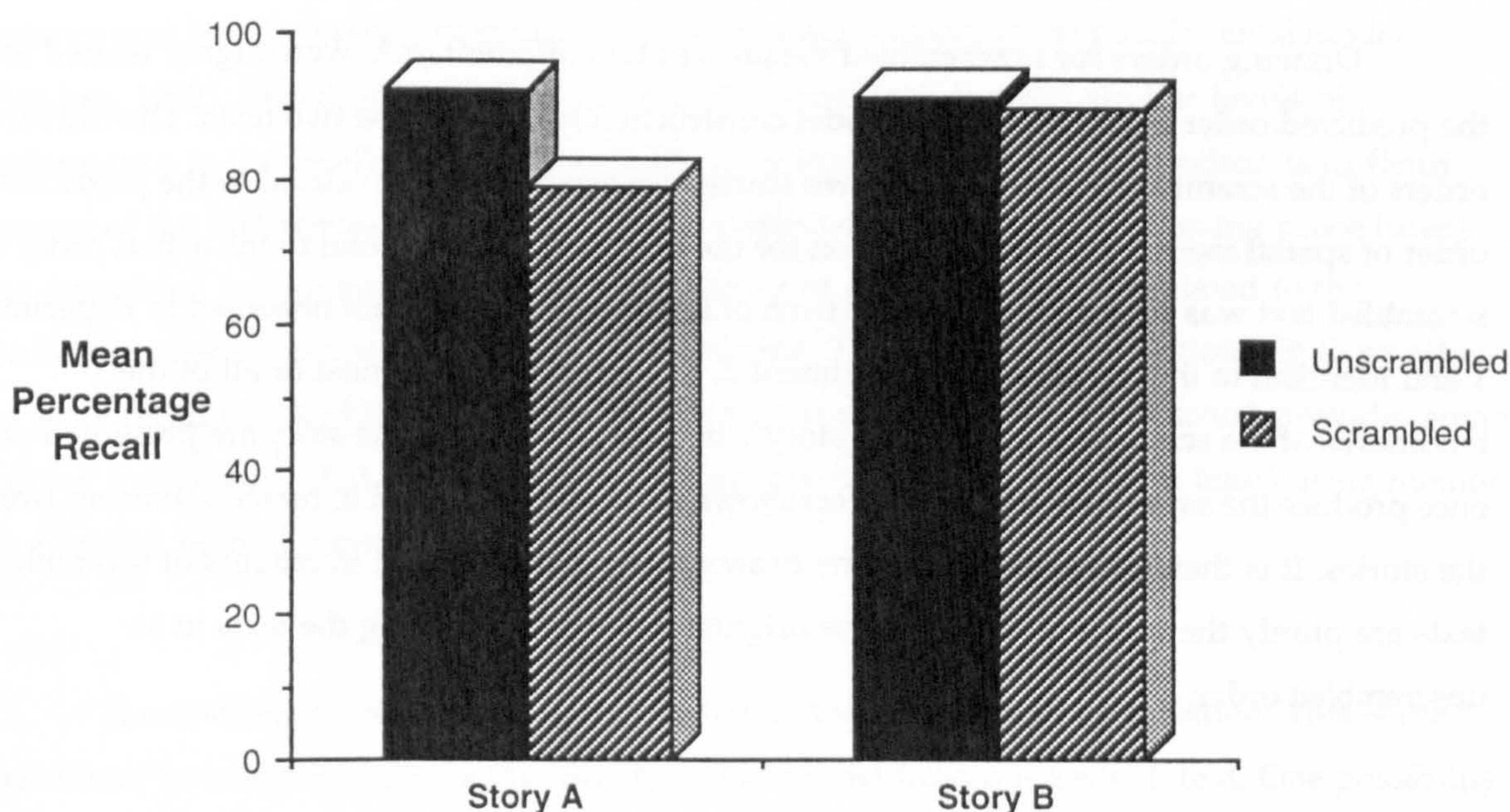


was confirmed with Fisher's Exact tests for both the unscrambled (two-tailed  $p = 0.0041$ ) and the scrambled text versions (two-tailed  $p = 0.027$ ).

### Levels of recall on the story recall task

The responses produced by subjects consisted of sentences or phrases containing two or more landmarks (usually, but not always linked by a spatial or temporal relation). Two scoring procedures were used to determine recall on the story recall task. The first scoring procedure used the criterion of exact verbatim recall; a sentence from the original text was scored as correctly recalled if it was reproduced exactly word for word. The second scoring procedure was as close as possible to the scoring of landmark recall in the drawing task. If a subject named or described a landmark during their attempt to rewrite the sentences of the original story that item was classified as correctly recalled. A name was considered correct if a part or all of it matched part or all of a single landmark from the story (e.g. "lake" for "Enchanted Lake"). A description was considered correct if it uniquely referred to the type or function of the landmark (e.g. "pub" for "White Horse Inn"). For example, a subject who wrote "One day princess Diana left her castle which was next to a magic lake" would be scored as incorrect for verbatim recall, but correct for recall of two landmarks (the 'Ancient Castle' and the 'Enchanted Lake').

Verbatim recall of the sentences of the original stories was at floor. Only one person was able to reproduce a sentence exactly. However, the scoring procedure was very strict and most subjects were able to produce some sentences that partially matched the original sentences they had read. Overall recall of landmarks was high. Figure 3.25 shows percentage recall averaged over all seven landmarks for each of the four possible texts versions (note that the relatively low recall for the scrambled version of story appears to be due to outliers).



**Figure 3.25: Mean percentage recall of landmarks in the story recall task for unscrambled and scrambled text versions of story A and B.**



Memory for landmarks in the story recall task was analyzed in the same fashion as that for the drawing recall task (using a trend test for dichotomous data). A significant relationship was obtained between the predicted order of spatial mental model construction and the probability of recall for each landmark ( $Z$  estimate = 3.46, one-tailed  $p < 0.0006$ ) for the scrambled version of story A. There was also a trend towards a negative relationship between landmark recall and the order that landmarks were mentioned in scrambled version of text A ( $Z$  estimate = -1.44, two-tailed  $p = 0.15$ ). Recall for story B and the unscrambled version of story A was almost at ceiling. No further tests were carried for these text versions because only a handful of subjects could be included in the analysis.

### 3.5.3 Discussion

This study replicated and extended findings from the two previous experiments. Two predictions were made for the drawing recall task. Firstly, it was predicted that, in general, drawing orders for both scrambled and unscrambled text versions should be similar to the order landmarks were mentioned in the original unscrambled text. However, where the initial spatial relation employed in a story is asymmetric (as in story A, but not story B) the landmark mentioned second should be the first item entered into a spatial mental model. So the spatial mental model account predicts that this landmark should be drawn first. For the story recall task two further predictions were made. Firstly, it was proposed that verbatim recall of the story should be poor. Secondly, it was proposed that there may be evidence that subjects reconstruct the text of the story from their memory for a spatial mental model.

Drawing orders for unscrambled versions of two different texts were highly related to the predicted order of spatial mental model construction for each of the two texts. The drawing orders of the scrambled versions of the two stories were more strongly related to the predicted order of spatial mental model construction for the unscrambled texts than to the actual order the scrambled text was presented in. This pattern of results is similar to that observed in Experiment 1 and identical to that obtained in Experiment 2. Readers who recall most or all of the landmarks of the scrambled version of a story when the sentences of the story are presented only once produce the same drawing order effect shown by people permitted to reread sentences from the stories. It is therefore unlikely that the drawing order effect found in readers of scrambled texts are purely the result of detecting the original order and rereading the texts in an unscrambled order.

In addition to predicting a relationship between overall drawing order and spatial mental model construction a more detailed prediction was made about the item that would be



drawn first. The spatial mental model construction process described in Chapter 2 makes a different prediction about the first two tokens entered into a spatial mental model according to the symmetry of the first spatial relationship in a description. Story A contained an asymmetric initial spatial relation. Story B contained a symmetric initial spatial relation. As predicted the first item drawn in story A was the second landmark mentioned in the description. For story B, however, the first item drawn was the first landmark mentioned in the description. This pattern of results was significant for both the unscrambled and scrambled versions of the two texts. The pattern is particularly significant because it appears to depend on the spatial characteristics of the relations used in the description. For this reason these results are difficult to explain without recourse to a representation constructed according to the spatial properties of the description.

Memory for landmarks in the drawing task replicated the pattern of recall obtained with the scrambled route text in Experiment 1. For the scrambled versions of both story A and story B there was a weak, but significant relationship between the predicted order of spatial mental model construction and the probability that an item was recalled. Neither story A nor story B showed any evidence that landmarks mentioned early on in the scrambled text were recalled better than those mentioned later.

Interpretation of the results of the patterns of recall on the story recall task is more problematic. Verbatim recall, as measured by the exact reproduction of the original sentences, was close to zero. However it would be premature to suggest that readers had no memory for the language of the text. A number of subjects produced partial reproductions of the original sentences. Unfortunately, it is impossible to rule out the explanation that partial recall of a sentence was based on reconstructing the sentence using memory for a spatial mental model (Fletcher, 1992). Recall of landmarks in the story recall task showed similar levels of performance to the recall of landmarks in the drawing task (detailed comparison is difficult because of the differences in the two tasks and potential differences in the scoring procedures). The scrambled text version of story A shows levels of recall significantly related to the predicted order of spatial mental model construction. The presence of a relationship in recall on the story recall task and the predicted order of spatial mental model construction provides some support for the view that the language of the text is reconstructed (in part at least) using memory for a spatial mental model.

The findings of this experiment do however raise a more serious question. This is the question of how readers construct a spatial mental model from a *scrambled* text. One possibility is that they try and remember the sentences of the story as they are presented and only when they believe they have reached the initial sentence from the *unscrambled* text do they begin to integrate the information contained in the sentences into a spatial mental model. A second



possibility is that readers begin the construction of a spatial mental model on reading the first sentence and integrate information from the remaining sentences as they are read. The former strategy results in a single, coherent spatial mental model constructed exactly in the order predicted by the account outlined in Chapter 2. It also places a heavy burden on a reader's working memory and language comprehension resources. The latter strategy reduces the burden of language comprehension but may require that two unrelated spatial mental models are constructed and retained until the relationship between the two models is established. This is because it is unlikely that *all* the early sentences will share a common landmark. In this case the *actual* order of spatial mental model construction will not be identical to the *predicted* order of spatial mental model construction. It should be noted, though, that the correlation between the actual and predicted order will probably still be quite high.

Perhaps the most likely strategy for readers to adopt is a hybrid one. Readers may defer construction of a spatial mental model until they have difficulty remembering the previous sentences or (more likely) until they establish a spatial relationship between landmarks from two different sentences. Such a hybrid strategy will still predict a spatial model construction order very close to that of the original, unscrambled text. This is because people will still have the overall goal of making sense of the situation described by the text. Most readers will find it easier to construct a spatial mental model once they that have identified the initial sentence of the unscrambled text (as in the first strategy). Even if they have already constructed a spatial mental model beginning with a different sentence (as is probable for the second strategy) information from the initial sentence is likely to reinforce the representation which has already been constructed. The main difficulty with all three strategies is that keeping track of all the information necessary to construct a spatial mental model is very demanding on working memory capacity. In this experiment this working memory burden can not be transferred to an 'external' memory resource because the sentences are removed from view once a subject has read them. This may explain why memory for landmarks from the last sentence of the scrambled texts is so poor (particularly in the case of story A).

Other factors may also influence drawing order. Some readers in Experiment 1 demonstrated a slight preference for a left to right drawing order, however, no indication of this was observed here (e.g. the sequence of landmarks for story A is predominantly right to left). The drawing orders produced by some readers may reflect the perceived temporal or causal order of the story rather than the order of spatial mental model construction. This could happen if the spatial mental model is elaborated by focusing on spatial relations according to their perceived temporal order. While this explanation appears plausible, it has difficulty accounting for all the findings reported here. For example, nearly every reader of story A drew the 'Enchanted Lake' first even though it was mentioned second. This is interesting because the 'Enchanted Lake' is not the starting point of the journey described in story A. This is difficult to account for in terms of perceived causal or temporal order. On the other hand it does make sense if it is considered to



be a reference point for another landmark and hence entered into a spatial mental model first. The patterns of recall for different landmarks suggest that (for some readers) certain landmarks are never fully integrated into a spatial mental model. If this is the case, it is difficult to explain solely in terms of the elaboration of spatial relations according to their perceived temporal order. This is because determining temporal order would require that most, if not all, the landmarks are present in the spatial mental model. It is more plausible that readers attempt to construct a representation using one of the strategies described above and that in most cases this results in a spatial mental model construction order close to that predicted in the unscrambled condition. However, the temporal and causal cohesion of the *unscrambled* route stories may explain why the relationship between spatial mental model construction order and drawing order is so strong when people recall an unscrambled text.

### 3.6 General Discussion

The results of the three experiments presented in this chapter have been used to support the view that the recall of spatial information from a description requires the retention of the order in which a spatial mental model is constructed. Three types of evidence have been considered. Firstly, evidence the drawing orders of subjects who have remembered most or all of the landmarks on being asked to draw the environment described by a short narrative. Secondly, evidence that the probability of a particular landmark being recalled depends on its location in the presented text or in the predicted order of spatial mental model construction. Lastly, evidence that the language of the text is considered as an alternative to the spatial mental model account.

The drawing orders of subjects in these studies replicate the effects first noticed by Taylor and Tversky (1992a). In every case drawing orders are strongly related to the predicted order of spatial mental model construction. Even where people read texts where sentences were presented in a scrambled order the same relationship between drawing order and the predicted order of spatial mental model construction was obtained. This relationship persisted even when each sentence was read only once, indicating that it could not simply be an artefact of rereading the text in an unscrambled order. In addition to replicating Taylor and Tversky's findings about overall drawing order, detailed predictions were made about which item people would draw first. Where an asymmetric spatial relation (e.g. 'to the left of') is used in the initial sentence of a description the second landmark in the sentence acts as a reference point for the first landmark. For this reason the spatial mental model construction process outlined in Chapter 2 predicts that the second landmark will be drawn first. Where the initial spatial relation in a description is symmetric (e.g. 'opposite') the meaning of the relation is not affected by the order of the two landmarks. In this case the spatial mental model construction order is likely to reflect the order the two landmarks are mentioned in the text. The analysis of drawing orders obtained in



Experiments 2 and 3 confirmed these predictions. These findings are consistent with the construction of a representation, such as a spatial mental model, that is constructed according to the meaning of the spatial characteristics of the relations used in the description.

The drawing orders observed in these experiments suggest that subjects are remembering or reconstructing a spatial mental model. Relatively little attention, however, has been addressed to the question of why people should utilize order information preserved in memory for the construction of a spatial mental model. New elements or tokens are entered into a spatial mental model in relation to existing tokens. As a consequence the integration of new information will tend to improve retention of existing tokens and relationships. The account therefore predicts that, all other things being equal, the initial tokens and relationships should be recalled better than later ones. Given that this order information is preserved in long-term memory why would this order be utilized during a drawing task? The main reason would be cognitive economy; adopting an existing order is far simpler than deciding on a new one. In particular, less demand would be placed on working memory resources. Morra and colleagues have provided evidence that working memory demands for drawing tasks, and especially for planning a drawing, are high (Morra et al., 1988). The tendency for the first few items to be less likely to deviate from the predicted drawing order suggests that working memory constraints may indeed be an important factor in the adoption of a particular drawing order. Once the initial items have been drawn, the planning load for the remaining items will have been reduced. This reduced working memory load may enable people to take account of other, more pragmatic, factors in the drawing task such as proximity. Analysis of the drawing orders in Experiments 2 and 3 suggests that many deviations from the predicted order occur in the middle portion of the route. Perhaps the consistency in the order in which the last items are drawn reflects pragmatic factors (the last items may be exits and so on) or a tendency for these items to be recalled only when all the other items are drawn. There is certainly evidence that the last few items in the predicted spatial mental model construction order are particularly hard to recall.

Where recall was not at ceiling it was possible to investigate the pattern of recall for different landmarks. A previous study by Denhière and Denis (1988) found evidence that the probability of recall of a given landmark from a spatial description was related to the order in which the landmark was integrated into a spatial representation. In Experiments 1 and 3 a significant relationship was obtained between the probability that an item was recalled from scrambled route texts and the predicted order of spatial mental model construction. A similar relationship was obtained between the probability of recall of items from the scrambled survey text in Experiment 1 and the order in which they were mentioned in that text. It was argued that this may be because readers are able to make partial sense of a scrambled survey text (and hence begin to construct a coherent spatial mental model) without successfully reordering it. While the observed effects are not as large as those for the drawing order analyses it is argued that they



reflect an advantage in retention for those landmarks that are most likely to be integrated into a spatial mental model early in its construction. These patterns of recall provide further support for the argument that the order that readers of scrambled texts construct representations of the situations described by the text is close (though not necessarily identical) to the predicted order of spatial mental model construction.

Evidence from the analyses of memory for the language of the text is less conclusive. In Experiment 2 subjects had good recognition memory for verbatim and paraphrase sentences. In Experiment 3 exact verbatim recall was poor, but memory for landmarks when rewriting the texts was close to ceiling. Some memory for the language of the text appears to be present, but it is not clear whether memory for the language of the text is sufficient for readers to reconstruct the spatial characteristics of the environment it describes. In Experiment 3 memory for the landmarks from the scrambled version of story A in the story recall task was significantly related to spatial mental model construction order. This result would favour the view that subjects were reconstructing the story based on their memory for the construction of a spatial mental model. Nevertheless it is very unlikely that subjects do not have some memory for the language of the text, however slight.

In this chapter a number of different arguments have been raised to support the view that recall on a drawing task is based on memory for a spatial mental model, or for the process of its construction. Many of these arguments refer to the way a particular experiment was carried out and have already been discussed in detail. Two remaining arguments which are of particular importance are reviewed briefly in the sections below.

### **3.6.1 Drawing orders reflect a spatial mental model rather than an image**

This argument was first put forward by Tversky (Taylor & Tversky, 1992a; Tversky, 1991). Tversky has argued that if people are drawing from a mental image then drawing order should reflect characteristics of the image rather than the process of its construction. For example if someone were given a copy of a map and asked to draw it they might begin by drawing its major structural features (e.g. boundaries, rivers, major roads and so on). If scanning a mental image shares similarities with scanning a real picture or map we would expect to find similar patterns in the drawing orders of someone who has learned a map. The results obtained by Kulhavy and his collaborators may reflect memory for an image (Kulhavy et al., 1993). However, many imagery theorists are sceptical about the notion of a distinct imagery code in long-term memory (Denis, 1991; Hampson & Morris, 1979; Kosslyn, 1980; Marschark, 1985). Also, as has already been pointed out, the lack of a relationship between order of presentation and the order or



probability of recalling a landmark in Kulhavy's studies may reflect the learning procedures he used or the levels of recall present in his experiments (see section 3.1).

### 3.6.2 Comprehending scrambled texts requires a mental model

Van Dijk and Kintsch (1983) provide an *a priori* argument why comprehension of a scrambled text is not possible without the construction of a representation of the situation described by the text (a situation or mental model). They argue that there are two ways to explain how people can understand scrambled stories:

"First, it may be that in retelling people reconstruct the story from the situation model they had formed when they heard it in scrambled form: As they had available the appropriate knowledge schemata, they were able to construct a canonical model from the scrambled input, and in retelling they simply work from that model rather than from the text representation proper. Alternatively, it might be the case that the representation itself was unscrambled and put into the correct order, in spite of the disorderly input. However, the only way this unscrambling could be done was to construct a situation model in canonical form, and then use this model to rearrange the textbase. In either case, the reordering of scrambled narratives presupposes the construction of a situation model."

[van Dijk and Kintsch (1983), p. 341]

Analysis of Van Dijk and Kintsch's argument suggests that in cases where people can bring to bear real-world knowledge (such as understanding a spatial description) they will concentrate on establishing a coherent 'situation model' rather than to go to the trouble of reorganizing their representation of the language of the text. This interpretation is at least partly borne out by the consistency of the relationship between spatial mental model construction order and drawing recall for scrambled texts.

## 3.7 Conclusions

The studies presented in this chapter have shown broad support for Tversky's position that drawing an environment described by a spatial description involves reconstructing a spatial mental model rather than recalling an image or a representation of the language of a text. A number of potential objections to this interpretation have been discussed and rejected. As always, though, there are a number of other issues that have not been dismissed.



Firstly, the material used in these studies (as well as those by Tversky) are relatively rich in semantic content; short texts relatively close to the kind of natural discourse such as verbal directions or descriptions of routes. While the studies may have gained in ecological validity they may have also lost in terms of generalizability. In particular, it is possible to argue that the findings are strongly influenced by the specific landmarks used in the stories or by the way the stories themselves were constructed. These criticisms are mitigated, to some extent, by the use of different stories and different types of story in each experiment.

Secondly, it has not been possible within this methodology to make a sharp distinction between memory for a spatial mental model and memory for its construction. It is possible that people only remember the process of spatial mental model construction (Payne, 1993). Evidence from studies by Tversky and others suggests that the structure of the mental model is also recalled (Glenberg & Langston, 1992; Tversky, 1991). Observation of the drawing orders from some subjects in the experiments presented here suggests that some of the deviations from the predicted drawing order may have been to insert a landmark in an intervening spatial location. If this is the case it would suggest that the structure of the spatial mental model is sometimes available.

Finally, it is not clear how well subjects in these studies remember the language of the text. Direct comparison of levels of recall for text memory and memory for a spatial mental model are difficult. Another factor is the possibility that verbatim recall may be influenced by guesswork strategies employed by subjects (Fletcher, 1992).

The experiments reported in Chapter 4 attempt to address all three of these issues. Less realistic and hence more strictly controlled materials are employed. Using methodology based on studies by Mani and Johnson-Laird (1982) and Payne (1993) it is possible to separate the influence of verbatim recall, memory for the structure of a spatial mental model and memory for the process of spatial mental model construction.



## Chapter 4 - Determinacy and memory for spatial descriptions

In a classic study of the distinction between propositional and analogical representations Mani and Johnson-Laird (1982) demonstrated that memory for the spatial layout generated from a description ('gist' recognition) is better for determinate descriptions than indeterminate descriptions. Mani and Johnson-Laird also found that memory for the sentences of the description ('verbatim' recognition) was better for indeterminate descriptions than for determinate descriptions (refer to section 2.4.5 for a summary of their experiment). The determinate spatial descriptions were consistent with only one arrangement of five objects. The indeterminate descriptions were consistent with two or more arrangements of the five objects. This 'crossover' interaction was interpreted as evidence that people construct spatial mental models to understand spatial descriptions. Determinate descriptions permit the construction of a single spatial mental model which supports memory for the structure or 'gist' of the situation. For indeterminate descriptions construction of a single consistent spatial mental model is not possible, therefore Mani and Johnson-Laird argued that people abandon their attempt to construct a spatial mental model and concentrate on remembering the description 'verbatim'.

Recently, Payne (1993) has proposed that people remember the operations used to construct a spatial mental model (an *episodic construction trace*) rather than the spatial mental model itself (see section 2.4.5 for a detailed discussion this hypothesis). Payne argued that recognition is dependent on the degree to which the process of spatial mental model construction overlaps between the original description and the to-be-recognized description (see Figure 4.1). In a direct test of the episodic construction trace hypothesis Payne showed that reordering the



sentences in the original descriptions depressed ‘verbatim’ and ‘gist’ recognition scores.<sup>1</sup> Reordering disrupts the episodic construction trace but has no effect on either the individual sentences or the situation they describe. Thus recognition could not depend on either memory for the sentences of the description or memory for the spatial mental model itself.

<u>Original description:</u>	<u>Operations in the construction process:</u>
The table is behind the stool	1. Enter stool into model
	2. Place table behind stool
The table is to the left of the bed	3. <i>Place bed to right of table</i>
The table is to the right of the chair	4. Place chair to left of table
The chair is behind the lamp	5. Place lamp in front of chair
<u>Reordered original description:</u>	<u>Operations in the construction process:</u>
The chair is behind the lamp	1. Enter lamp into model
	2. Place chair behind lamp
The table is to the right of the chair	3. Place table to right of chair
The table is behind the stool	4. Place stool in front of table
The table is to the left of the bed	5. <i>Place bed to right of table</i>

**Figure 4.1: The operations used to construct a spatial mental model from a reordered description.**

Two of the studies presented in this chapter (Experiments 4 and 6b) investigate the hypothesis that memory for spatial relationships includes both components. Initially people may focus on the construction of the spatial mental model and the process of construction is recorded in long-term memory. Once the spatial mental model has been constructed the representation is elaborated and the structure of the spatial mental model is recorded in long-term memory. These would correspond to the *construction* and *consultation* stages of spatial mental modeling (see section 2.5.1). In Experiment 4 the number of times subjects are presented with each description (once or three times) is manipulated. It is predicted that with only one presentation people will tend to remember only the episodic construction trace. With three

---

<sup>1</sup> It is important to note the distinction between ‘gist’ and ‘verbatim’ recognition *scores* and gist and verbatim memory. Payne (1993) has argued that it is possible to get high scores on ‘gist’ or ‘verbatim’ recognition test with no memory whatsoever for the language of the description or for the spatial relationships present in a spatial mental model. This is because under most circumstances the inferable description has a greater overlap with the original description than either of the two foils.



presentations people are also likely to remember the spatial mental model itself. Overall, the pattern of data should provide evidence for both the episodic construction trace *and* memory for the structure of a spatial mental model.

Evidence of memory for an episodic construction trace is obtained if reordering descriptions in a recognition test reduces recognition memory. Reordering the original description changes the nature of the operations used to construct a spatial mental model. This is because the operations carried out to construct the spatial mental model depend on the tokens which are already present in the spatial mental model (this is demonstrated in Figure 4.1). However, there is no reason to expect reordering to influence memory for the sentences themselves, because each individual sentence is intact. In addition, the reordered descriptions still describe exactly the same arrangement of objects, so reordering should not influence memory for the structure of a spatial mental model. Memory for the structure of a spatial mental model would be indicated by the crossover interaction predicted by Mani and Johnson-Laird. This would be indicated by high 'gist' recognition for the determinate descriptions and high 'verbatim' recognition scores for the indeterminate descriptions. Finally, memory for the sentences of the descriptions can be ascertained by looking at 'verbatim' recognition scores for the reordered descriptions.

The remaining experiment presented in this chapter (Experiment 5) takes the opportunity to re-examine some of the issues raised by the experiments in Chapter 3. This is done by investigating memory for the spatial descriptions presented in Experiment 4 using a drawing task. Thus it is possible to look at drawing orders and the probability of recall for descriptions where there is no narrative structure to influence recall. Also by looking at eight different descriptions the influences on recall of individual items or landmarks (which may be more or less salient) is minimized.



# 4.1 Experiment 4

In this experiment recognition memory and recall for determinate and indeterminate spatial descriptions is investigated. Evidence will be sought for and against three ways of remembering spatial mental models; an episodic construction trace, a verbatim (or propositional) trace of the language of the description or memory for the structure of a spatial mental model. The experiment is based on those of Mani and Johnson-Laird (1982) and Payne (1993). In phase 1 of the experiment subjects learn determinate and indeterminate verbal descriptions for a simple diagram matching task. This task involves deciding whether a diagram presented by the experimenter is consistent with the original description or not (see Figure 4.2). The purpose of the diagram matching task is simply to ensure that subjects read and learn the spatial descriptions (they are not informed that there will be any further tests of their memory for the descriptions). In the second phase of the experiment subjects are given a surprise recognition test. Subjects are presented with four possible 'target' descriptions (the original description, an inferable description describing the same spatial configuration of objects and two foils). Their task is to rank the descriptions 1, 2, 3, or 4 according to how closely they resemble the original description (a description they were presented with in phase 1). In phase 3 subjects are asked to write down how they carried out the surprise recognition test and what they were thinking while they performed the task.

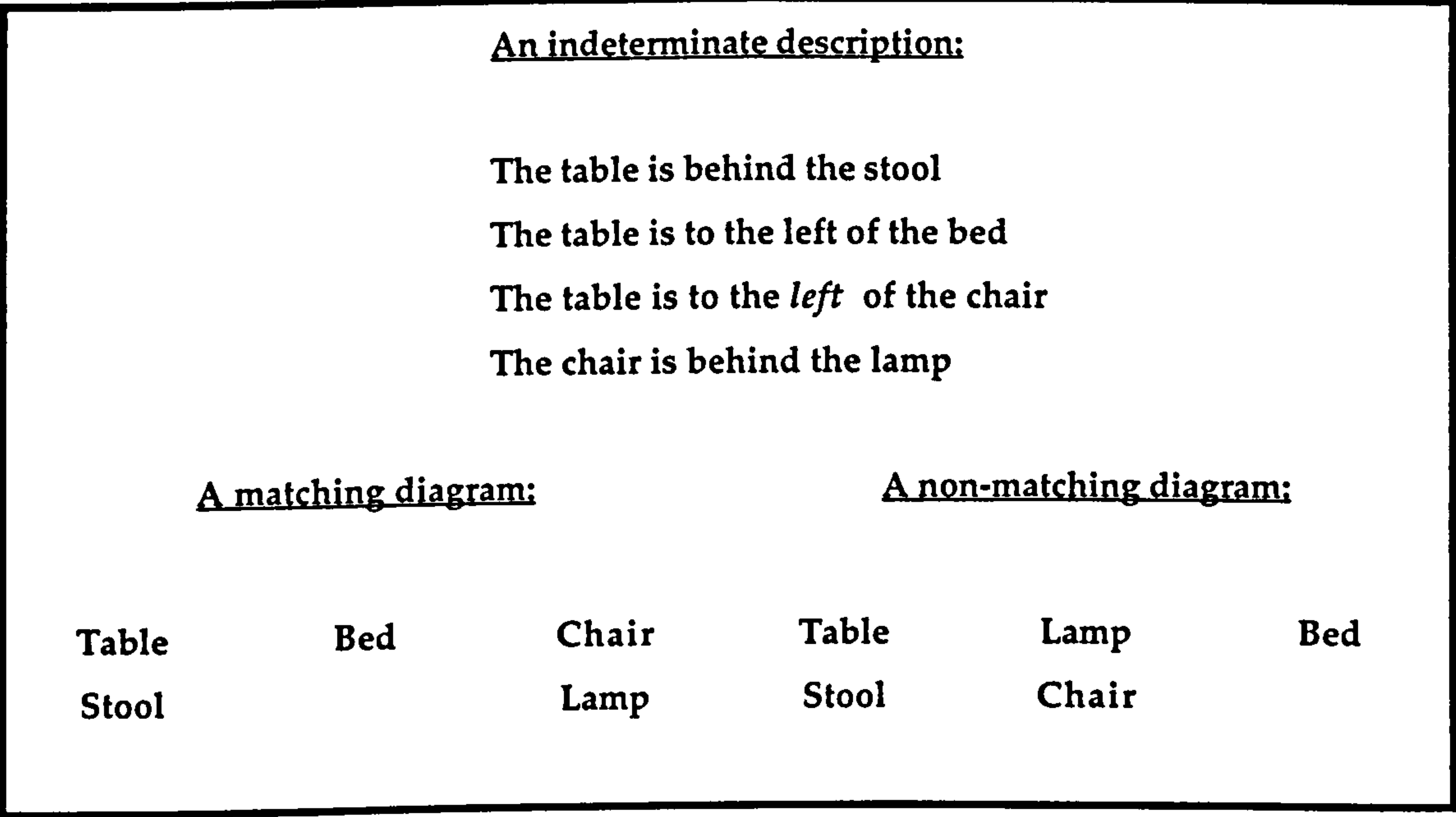


Figure 4.2: An example of a matching and non-matching diagram for an indeterminate description.

This experiment differs from that performed by Payne or Mani and Johnson-Laird in several details. Firstly, two groups of subjects were used. The first group of subjects read and matched each description in phase 1 only once (the *one diagram* condition). The second group of



subjects read each description three times, matched each description to three different diagrams and also received feedback (the *three diagram* condition).

<u>Original indeterminate description:</u>	<u>Corresponding episodic construction trace:</u>
The duck is behind the cat	[start duck cat behind]
The cat is to the left of the frog	[frog cat right]
The duck is behind the mouse	[mouse duck front [clash mouse duck cat]]
The wolf is to the right of the duck	[wolf duck right]
 <u>Reordered description:</u>	
<i>The wolf is to the right of the duck</i>	[start wolf duck right]
<i>The duck is behind the mouse</i>	[mouse duck front]
<i>The duck is behind the cat</i>	[cat duck behind [clash cat duck mouse]]
<i>The cat is to the left of the frog</i>	[frog cat right]
 <u>Inferable description:</u>	
The wolf is behind the frog	[start wolf frog behind]
<i>The cat is to the left of the frog</i>	[cat frog left]
The duck is to the left of the wolf	[duck wolf left]
The mouse is in front of the duck	[mouse duck front [clash mouse duck cat]]
 <u>Foil (created from original description):</u>	
The mouse is behind the cat	[start mouse cat behind]
<i>The cat is to the left of the frog</i>	[frog cat right]
The mouse is behind the wolf	[wolf mouse front [clash mouse wolf cat]]
The duck is to the right of the mouse	[duck mouse right]
 <u>Foil (created from inferable description):</u>	
The duck is behind the frog	[start duck frog behind]
<i>The cat is to the left of the frog</i>	[frog cat right]
The mouse is to the left of the duck	[mouse duck left]
The wolf is in front of the mouse	[wolf mouse front [clash wolf mouse cat]]

Figure 4.3: Examples of the reordered, inferable and foil descriptions used in the surprise recognition test (sentences or propositions shared with the original description are in italics).



The purpose of this manipulation was to induce greater elaboration or a greater amount of learning of the original descriptions. For the surprise recognition test (phase 2) half the original descriptions were reordered. In addition, all the descriptions apart from the stable originals (i.e. the original descriptions which were not reordered) were constructed to have an episodic construction trace overlap of one with the original description. Figure 4.3 shows how each alternative shared only one proposition of its episodic construction trace with the original description. Each description in the surprise recognition test (apart from the original description) also shared only one sentence with original description. In Figure 4.3 shared sentences and shared propositions from an episodic construction trace are shown in italics. This manipulation means that memory for an episodic construction trace can only support verbatim recognition of the stable original description (assigning the stable original description the rank '1'). Similarly, memory for the sentences will support verbatim recognition of both the stable and reordered original description (assigning the stable or reordered original description the rank '1'). Neither an episodic construction trace, nor memory for the sentences can support 'gist' recognition in the absence of verbatim recognition (assigning the inferable description the rank '1'). High levels of 'gist' recognition (for example for the determinate descriptions) could only be supported by memory for the structure of a spatial mental model.

The predictions made by three different kinds of memory for spatial descriptions are summarized in Figure 4.4:

<u>What is remembered:</u>	<u>Prediction:</u>
The language of the description (comprehension processes)	i) No effect of reordering. ii) Verbatim recognition higher than gist.
An episodic construction trace (construction processes)	i) Reordering reduces recognition. ii) Verbatim recognition higher than gist.
The structure of a spatial mental model (consultation processes)	i) High determinate gist recognition. ii) High indeterminate verbatim recognition.

**Figure 4.4: The predictions from three different accounts of memory for spatial descriptions.**



If people only remember the language of a description reordering should have no effect on verbatim recognition scores. Also, because only the original description contains more than one sentence from the original description it also predicts that gist recognition scores (based on a high rank for the inferable description) can not exceed verbatim recognition scores. The episodic construction trace account predicts that reordering the original description should depress verbatim recognition scores to chance levels in phase 2. The account also predicts that gist recognition scores will never exceed verbatim recognition scores, because the inferable description always has a trace overlap of one. Finally, the classic mental models account (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982) predicts that verbatim recognition scores should exceed gist recognition scores for the *indeterminate* descriptions and that gist recognition scores should exceed verbatim recognition scores for the *determinate* descriptions. In fact, support for the mental models account would be obtained provided gist recognition scores exceed verbatim recognition scores for the determinate descriptions. Though not as strong a prediction as that from the classic mental models account, such a result could only be explained by memory for the spatial relationships present in the descriptions.

#### 4.1.1 Method

##### Subjects

Thirty-six people volunteered to participate in the experiment (most were graduate students or staff from the Open University). Subjects were offered three pounds for taking part in the experiment. The ages of the subjects ranged from early twenties to early sixties.

##### Design

Three tasks were used in the experiment; a diagram matching task (phase 1), a surprise recognition task (phase 2) and a retrospective protocol task (phase 3). Subjects were randomly allocated to either the *one diagram* condition (where each description and was presented once before a single diagram in the matching task) or the *three diagram* condition (where each description was presented once before each of three diagrams). The three within subject factors were the determinacy of the spatial description (*determinate* or *indeterminate*) whether more matching diagrams than non-matching diagrams were seen by subjects in phase 1 (*matching* or *non-matching*) and whether the spatial description was reordered in phase 2 (*stable* or *reordered*). Two dependent variables were used for the analysis. These were *verbatim* recognition and *gist* recognition (see the results section for details of how these were scored). The eight spatial descriptions were randomly assigned to these eight conditions.



## Materials

Descriptions for phase 1 were based on the recipes of Payne (1993). Four determinate and four indeterminate descriptions were used (excluding the two practice items), each containing four sentences and describing five objects. Referential continuity was maintained by introducing one new object in each sentence after the first. Determinate and indeterminate descriptions differed only by a change of one word (e.g. 'left' instead of 'right'). Each description was headed by the category of objects described (e.g. Creatures). Categories were chosen to minimize confusion between descriptions, so that each of the eight categories was associated with a single condition of the factorial design. Each description described an F or C shaped array or a rotation or mirror image of such an array. For each description up to three diagrams were produced for the matching task. The diagrams depicted either a permissible configuration of the objects in the description (a matching diagram) or a configuration where only two of the four spatial relations in the description were satisfied (a non-matching diagram).

For phase 2 a set of four descriptions were created for each description category in phase 1. For each set only one sentence was shared between the four descriptions. One description was the original, either in a stable form (with a trace overlap of 4) or in a reordered form (with a trace overlap of 1). A second description was inferable from the original, it described the same array of five objects but retained only one sentence from the original. The inferable description was constructed by reversing two relations ('x is to the right of y' in place of 'y is to the left of x') and adding a new relation true of (but not explicit in) the original. Two foils were created by interchanging three of the five original object names for the inferable and original descriptions (the two foils thus describe the same configuration of objects). These changes meant that only one sentence from the original description was left unchanged. Each foil retained a trace overlap of one with the original description as well as preserving its shape and general structure. For examples of the reordered, inferable and foil descriptions for one of the descriptions used in the experiment refer back to Figure 4.3 above.

## Procedure

The experiment was in three phases. At the start of phase 1 (the diagram matching task) subjects were presented with instructions for the matching task and given two practice trials (one determinate and one indeterminate). The purpose of the phase 1 task was to ensure that subjects made an attempt to learn the descriptions before proceeding to phase 2 (however, note that subjects were not informed about phase 2 prior to carrying out the surprise recognition test). The practice descriptions were presented on cards, which subjects read until they indicated they were ready. At this point the description was removed and a diagram presented. Subjects were asked to circle 'good match' or 'bad match' on a prepared scoring sheet according to whether the diagram was consistent with the original description. For the two practice trials feedback (the word 'correct' or 'incorrect') was given by the experimenter. Following the practice trials subjects



were encouraged to ask questions about the task. After the practice items the eight test items (two determinate matching, two determinate non-matching, two indeterminate matching and two indeterminate non-matching) were presented in a random order. The procedure for the eight test items differed between the one and three diagram conditions. In the *one diagram* condition each test item was followed by a single diagram before proceeding to the next item. No feedback was given. In the *three diagram* condition each test item was followed by three diagrams. Before each new diagram the subject was re-presented with the original description and asked to reread it at least once. After each of these diagrams feedback was given as in the practice trials. For the matching descriptions two of three diagrams were good matches, for the non-matching condition only one was a good match. Before and after phase 1 there were no differences between the one and three diagram conditions.

After phase 1 all subjects were given written instructions for phase 2 (the surprise recognition test). In phase 2 subjects were asked to rank four alternatives for each category of description presented in phase 1. The alternatives consisted of written descriptions similar to those presented in phase 1 (see Figure 4.3). Ranks were assigned according to how closely each alternative resembled the original description the subject had read. The four alternatives labelled A, B, C and D were presented in a random order on a single sheet of paper. A different size and style of font was used from that in phase 1. Subjects were told that in every case one of the four alternatives contained the four sentences of the original description, though not necessarily in the original order. Subjects were asked to rank this alternative 1 and the other three alternatives 2, 3 and 4 according to how closely they resembled the original description. They were asked to rank all four alternatives even if they were certain of their first choice (an example of the recognition test materials for one of the eight descriptions is given in Figure 4.3).

Phase 3 followed on immediately after the recognition task in phase 2. Subjects were given a blank sheet of paper and asked to briefly note down how they had performed the previous task and in particular what they were thinking as they performed it. For a sample of the 36 subjects a backward digit span test and the Daneman and Carpenter reading span test were carried out after completing the experiment (Daneman & Carpenter, 1980).

#### 4.1.2 Results

##### Performance on the diagram matching task (phase 1)

Overall performance on the diagram matching task was high (86%). All thirty-six subjects scored over fifty percent for both the determinate and indeterminate descriptions, thus indicating that performance was reliably above chance.



Performance on the recognition task (phase 2)

Scoring

Two different scoring procedures were employed. The first procedure was that used by Payne (1993) and was used to calculate percentage recognition scores for each subject.

<u>Description:</u>	<u>Ranking A</u>	<u>Ranking B</u>	<u>Ranking C</u>	<u>Ranking D</u>
Original	1	2	1	4
Inferable	2	1	4	1
Foil A	3	4	2	2
Foil B	4	3	3	3
<u>Type of recognition scored:</u>				
Verbatim	Yes	No	Yes	No
Gist	Yes	Yes	No	No

Figure 4.5: Four examples of the scoring procedure used by Payne (1993).

Successful verbatim memory was scored when subjects ranked the original description (stable or reordered) first. Successful gist memory was scored when subjects ranked both the inferable and the original description (in either order) higher than the two foils. Examples of how four rankings would be classified is provided in Figure 4.5. The percentage recognition scores are used in the summary statistics, tables and graphs later in the results section. The percentage recognition scores are also used to make pairwise comparisons within subjects (using the Wilcoxon signed ranks test) or between subjects (using the Mann-Whitney U test).

A second scoring procedure was adopted in order to test for the crossover interaction predicted by Mani and Johnson-Laird. A separate scoring procedure was necessary to analyze this interaction using a chi square design. The chi square test requires that the variables being manipulated are independent, so for this scoring procedure subject responses which exhibit both verbatim and gist recognition procedure are not included in the analysis. In effect this means that only one type of response was not included (this type of response is indicated by 'Ranking A' in Figure 4.5). However as they represent tied recognition scores they do not contribute to the results of the pairwise comparisons reported in the results section. One additional difference between the two scoring procedures was prompted by the possibility that the scoring method used by Payne underestimates the actual level of gist recognition. Inspection of Figure 4.5 will show that in 'Ranking D' the inferable description is ranked '1' but gist recognition is not scored (because the



subject has not ranked the original description above the two foils). However, if some subjects assume that only one description describes the original arrangement of objects, they may rank the inferable '1' and rank the remaining three descriptions using some other criterion. For the chi square analysis gist recognition is scored if the inferable description is ranked first. Note that if, Payne's scoring procedure is more accurate, the worst outcome from adopting this scoring procedure is that gist recognition scores will be slightly inflated by chance in *all* the conditions. Therefore there is no reason to believe that this new scoring procedure biases the results in any direction.

*Between group differences*

Differences between the one diagram and three diagram groups were analysed using Mann-Whitney U-tests. No significant differences were found between groups on verbatim and gist scores in phase 2 for any of the four matching descriptions or on the matching task in phase 1. For the subsequent analyses data were pooled across both groups of subjects.

*Overall recognition performance*

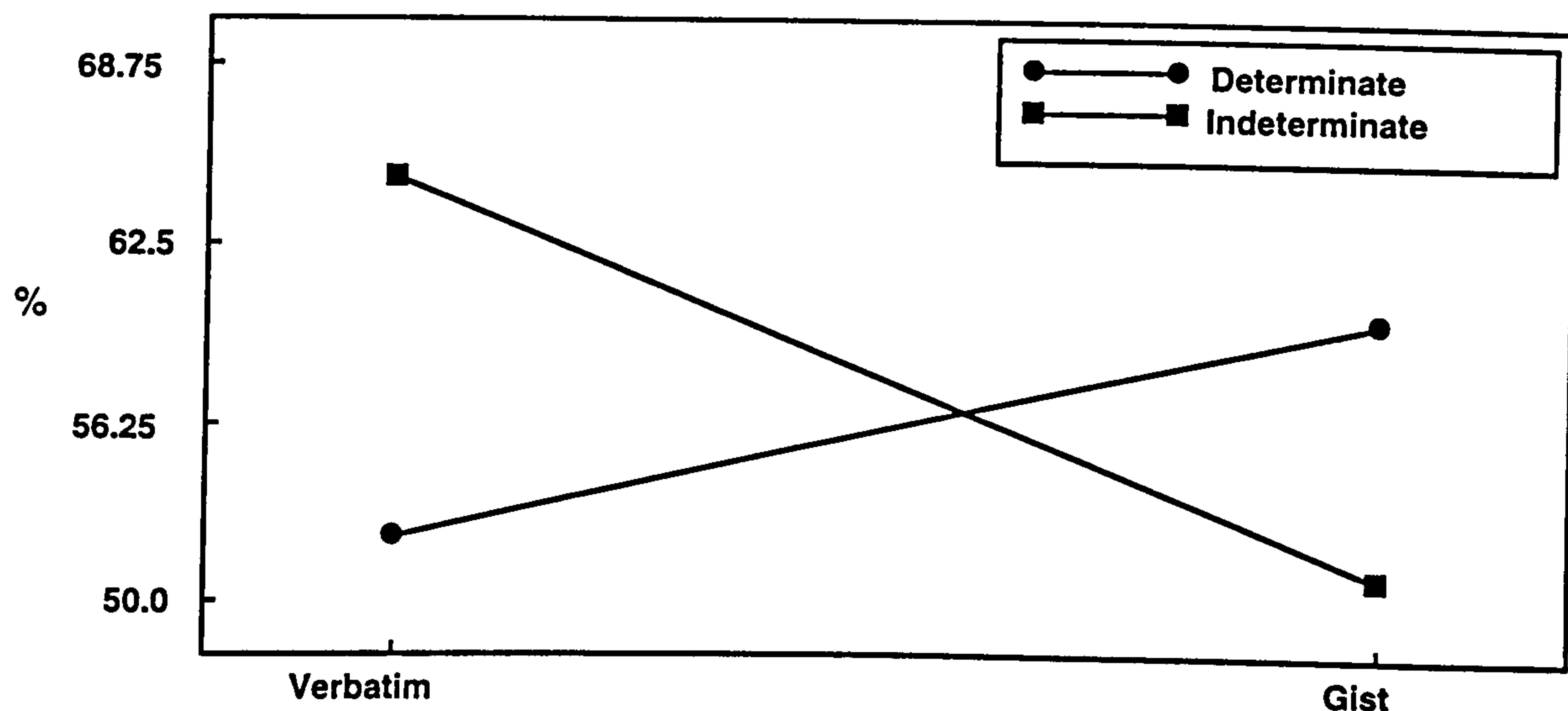
Chance performance is 25% for verbatim and 16.7% for gist recognition. Mean recognition scores are summarized by description type in Figure 4.6 below:

<u>Description type</u>		<u>Verbatim</u>	<u>Gist</u>
Determinate	Stable	56.9 %	55.6 %
	Reordered	50.0 %	62.5 %
Indeterminate	Stable	70.8 %	48.6 %
	Reordered	59.7 %	55.6 %

Figure 4.6: Mean percentage verbatim and gist recognition by description type

Overall recognition performance (averaged over stable and reordered descriptions) showed a pattern consistent with memory for a spatial mental model; gist memory was higher for determinate descriptions and verbatim memory was higher for indeterminate descriptions. Figure 4.7 shows the interaction of description type (determinate or indeterminate) type of recall (gist or verbatim). The significance of this interaction was tested using a chi square design. Rankings which exhibited both gist *and* verbatim recognition were excluded from the analysis (refer to the section on scoring above). The analysis thus compared the frequency with which determinate and indeterminate descriptions supported gist *or* verbatim recognition (data points were pooled across subjects as well as across descriptions).





**Figure 4.7: Mean percentage gist and verbatim recognition by type of description.**

The contingency table for the chi square test is shown in Figure 4.8 below:

		Type of recognition:	
		Verbatim	Gist
Type of description:	Determinate	23	34
	Indeterminate	38	18

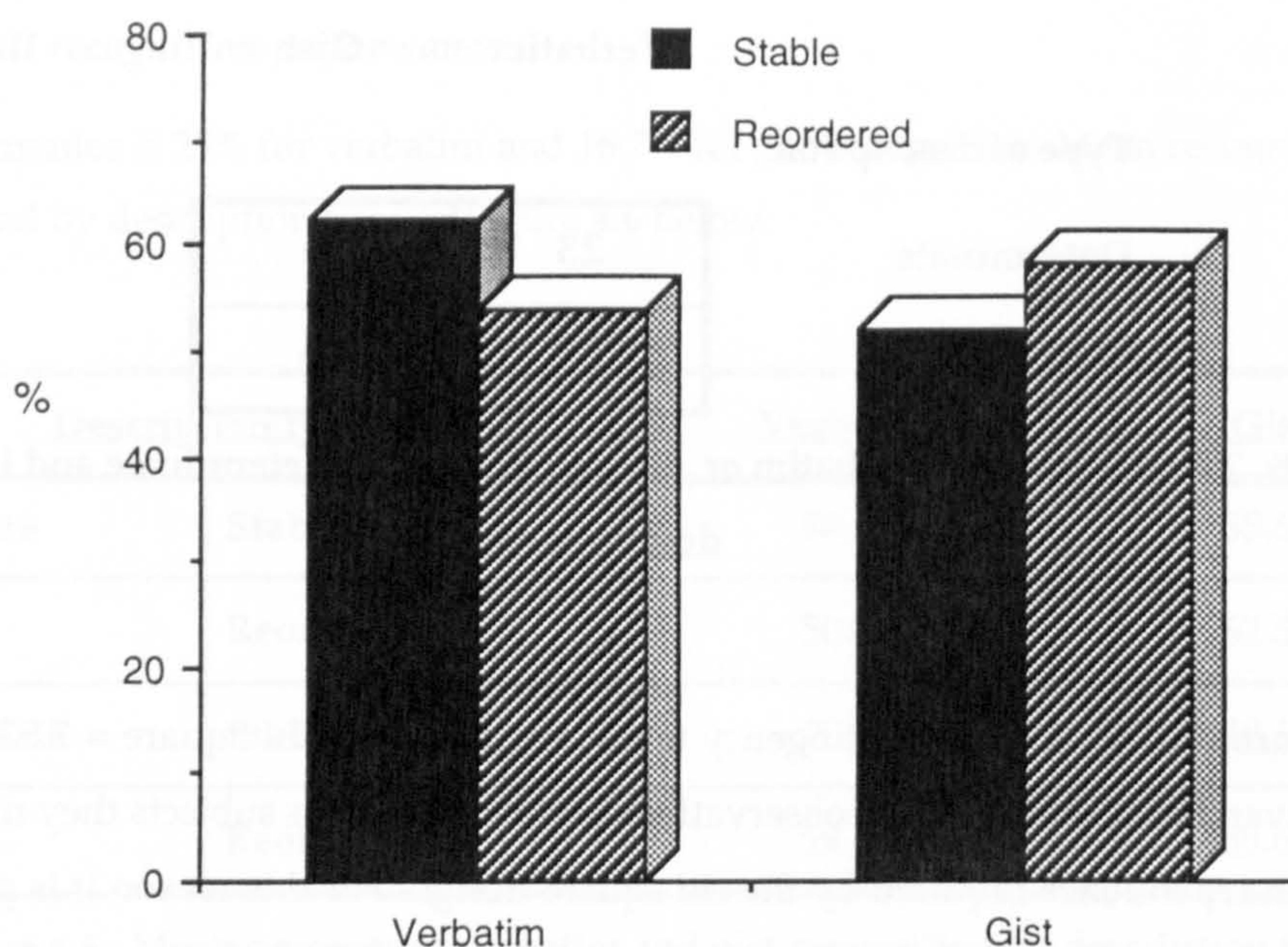
**Figure 4.8: The frequency of verbatim or gist recognition for determinate and indeterminate descriptions.**

The chi square statistic for this contingency table is significant (Chi Square = 7.53, d.f. = 1,  $p < 0.01$ ). However, because individual observations are pooled across subjects they may not meet the criteria for independence required by the chi square design. For this reason it is probable that the value of the chi square statistic has been inflated. A more conservative test of this prediction is to make two pairwise comparisons using the Wilcoxon signed-ranks test (this was the method of analysis used by Mani and Johnson-Laird). The first comparison is between verbatim recognition scores for the determinate and indeterminate description ( $N = 36$ ,  $Z = 2.25$ , one-tailed  $p = 0.012$ ). The second comparison is between gist recognition scores for the determinate and indeterminate description ( $N = 36$ ,  $Z = 1.85$ , one-tailed  $p = 0.032$ ). Even adopting this conservative method of testing this prediction it is clear the results support the spatial mental model account. People show significantly higher gist recognition for the determinate descriptions and significantly higher verbatim recognition for the indeterminate descriptions.



The effect of reordering on memory was analyzed separately for gist and verbatim recognition. Figure 4.9 shows mean percentage gist and verbatim recognition for stable and reordered descriptions. Differences were tested using the Wilcoxon test. While verbatim recognition for stable descriptions was greater than that for reordered descriptions this did not reach significance ( $N = 36$ ,  $Z = 1.37$ , two tailed  $p = 0.17$ ). For gist recognition the effect of reordering was reversed, though this too did not reach significance ( $N = 36$ ,  $Z = -1.68$ , two-tailed  $p = 0.093$ ).

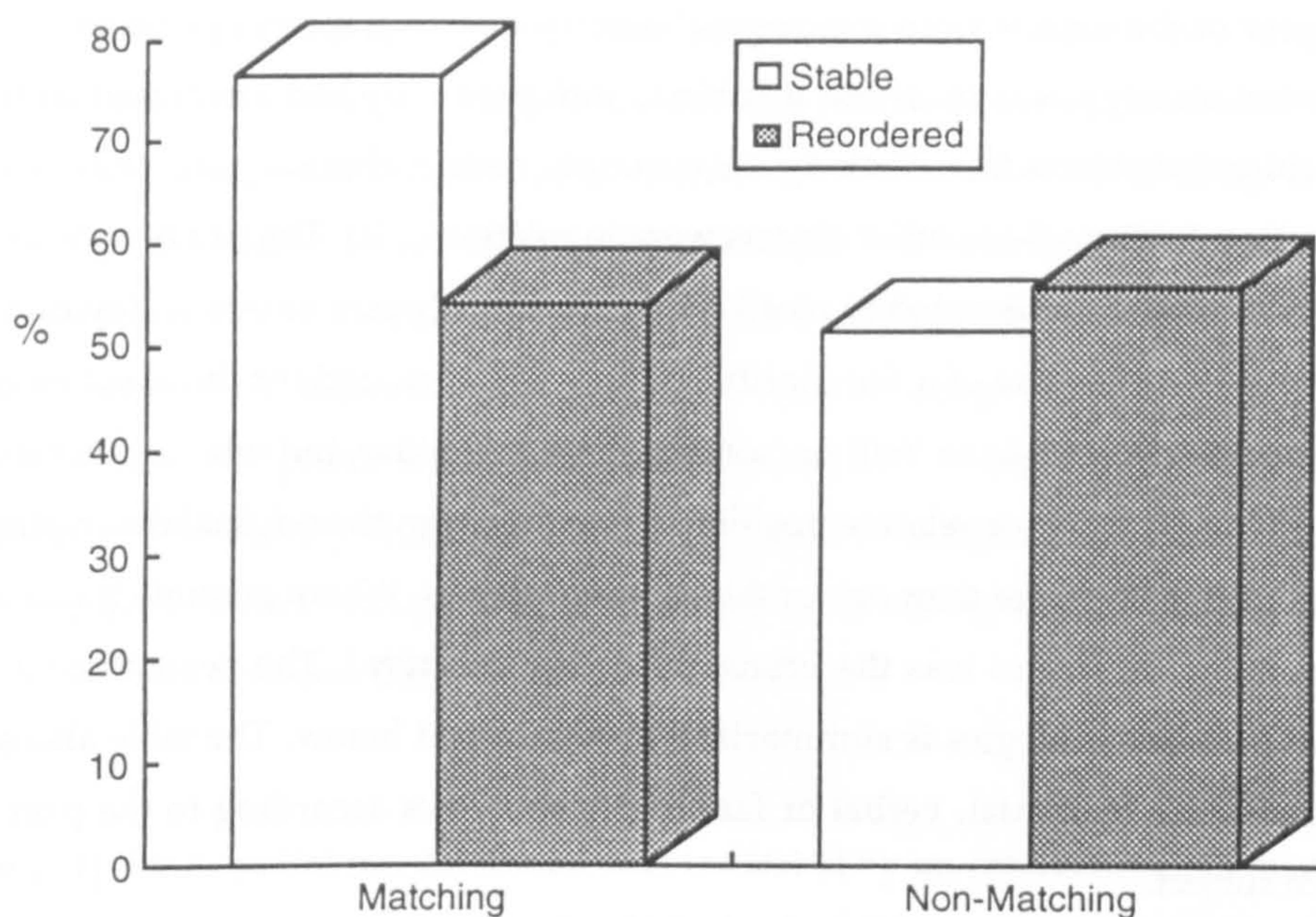
The effects of reordering were more marked for some descriptions than others. During the analysis it was noticed that the effect was more consistent for the matching descriptions than for the non-matching descriptions. This finding is difficult to explain in terms of the number of matching diagrams subjects were exposed to in phase 1 (the only consistent difference between the matching and non-matching condition). The matching/non-matching variable also has no influence on the construction of the recognition test materials. The findings were also stable across the one and three diagram conditions. This is particularly strange as the subjects in these two conditions saw different diagrams (and of course different numbers of diagrams).



**Figure 4.9: Mean percentage gist and verbatim recognition for stable and reordered descriptions.**

A more plausible explanation is that the results were due to differences in difficulty between descriptions which were coincidentally more pronounced among the descriptions used in the non-matching conditions. This is borne out by the variation in levels of recall for the non-matching descriptions in Experiment 5 which used the same descriptions and involved some of the subjects as Experiment 4 (see Experiment 5 below). Figure 4.10 shows levels of verbatim recognition for the matching and non-matching descriptions.





**Figure 4.10: Mean percentage verbatim recognition for stable and reordered descriptions.**

Verbatim recognition on the matching descriptions is significantly impaired by reordering ( $N = 36$ ,  $Z = 3.02$ , two-tailed  $p = 0.0025$ ). For the non-matching descriptions verbatim recognition for the reordered description is higher than that for the stable descriptions (though this effect is not significant). Gist recognition for the matching and non-matching descriptions follows the same pattern with matching gist recognition also being significantly impaired by reordering ( $N = 36$ ,  $Z = 2.63$ , two-tailed  $p = 0.0086$ ).

In addition to comparing different levels of verbatim and gist recognition the level of association between the two variables was also examined. Verbatim and gist recognition were correlated using Spearman's Rho. A significant relationship between gist and verbatim recognition was obtained ( $Rho = 0.431$ ,  $N = 36$ , two-tailed  $p = 0.011$ ). This indicates that verbatim recognition may form the basis of gist recognition for some subjects (though it could also indicate that verbatim recognition is a by-product of gist recognition).

### Individual differences

#### *Retrospective protocols*

Written reports made by subjects immediately after the recognition task were classified according to a several pre-defined categories; visuo-spatial strategy, spatial strategy, verbal strategy or familiarity strategy. The term strategy is used here to include both explicit recognition strategies as well as a preference for certain types of information for the purposes of establishing recognition. The use of a visuo-spatial strategy was recorded when subjects reported



visual imagery of the objects from the original descriptions (e.g. mental pictures or images).<sup>2</sup> The use of a spatial strategy was recorded if subjects indicated they had attempted to recall the relative positions of objects from the original descriptions (e.g. choosing one object as a reference point and remembering where other objects were in relation to it). The use of a verbal strategy was recorded if subjects attempted to recall sentences, key phrases or whole descriptions from the original description. The use of a familiarity strategy was included for those subjects who reported that some descriptions ‘felt’ or ‘sounded’ more familiar, but who did not explicitly report recognizing the text or relative positions of objects from the original descriptions. Some subjects reported using more than one of the above strategies. Where possible it was noted which if any of the above strategies was the primary strategy described. The percentage distribution of the different primary strategies is summarized in Figure 4.11 below. The table also shows the presence or absence of spatial, verbal or familiarity strategies according to the primary strategy used by each subject.

	Primary Strategy	Spatial Strategy present	Verbal Strategy present	Familiarity Strategy present
Visuo-spatial	28%	100%	50%	40%
Spatial	33%	100%	42%	33%
Verbal	8%	33%	100%	0%
Familiarity	14%	100%	0%	100%
Other/Unclear	17%	67%	67%	17%

**Figure 4.11: The distribution of different primary and secondary recognition strategies**

Data from subjects reporting primary spatial or visuo-spatial strategies were analyzed further. There were no significant differences in recognition scores between subjects reporting spatial strategies and those reporting visuo-spatial strategies. Direct comparison of spatial or visuo-spatial strategies with other strategies was not possible because of the small group sizes.

*The relationship between protocols and recognition performance*

For each subject in the experiment recognition performance was classified as being consistent with memory for the structure of a spatial mental model or memory for an episodic construction trace. Each subjects’ performance was scored as ‘for’, ‘against’ or ‘neutral’ with respect to each

---

<sup>2</sup> One subject reported visual imagery of the text of the description itself; this was classified as a verbal strategy.



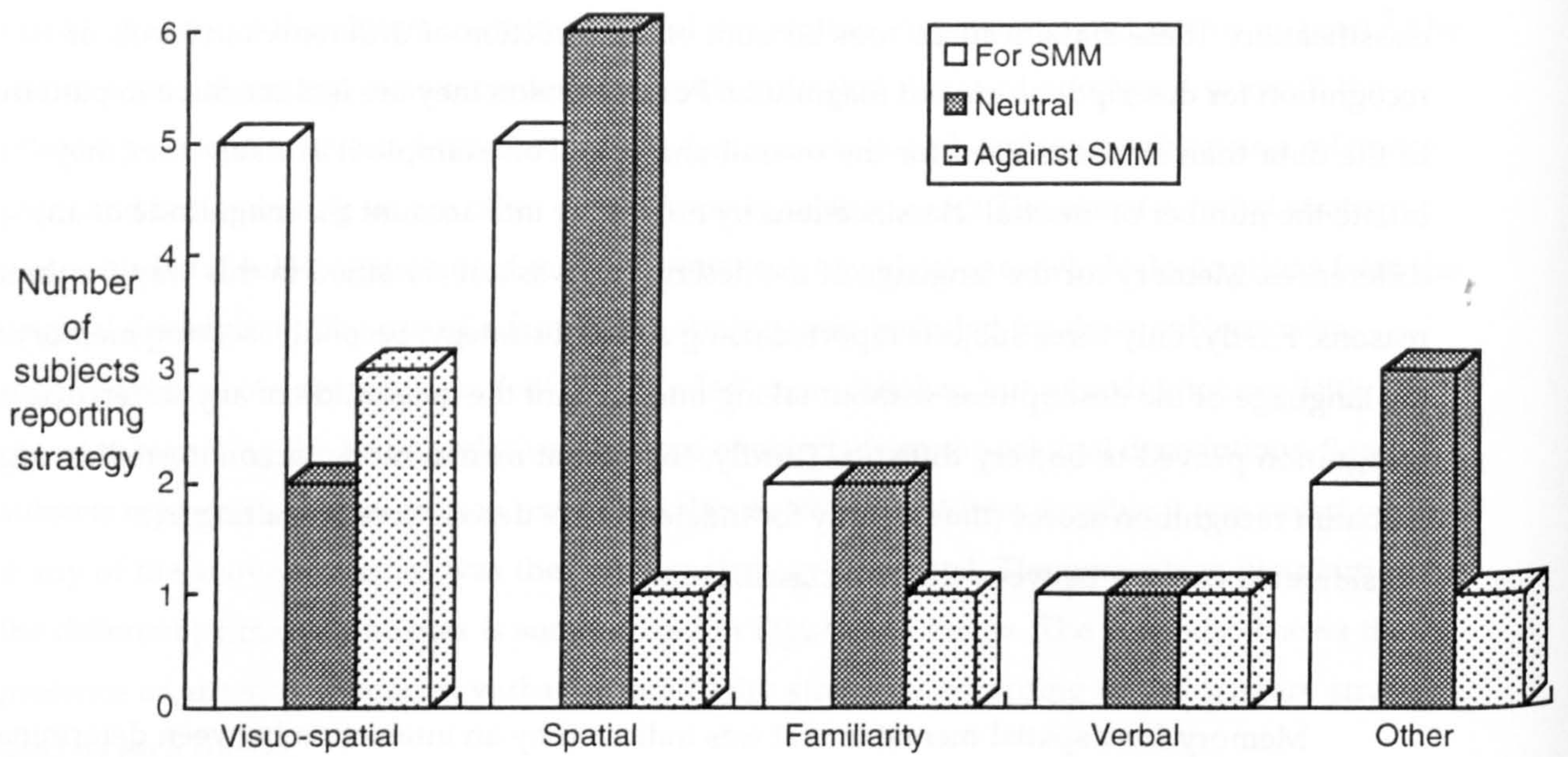
classification. These classifications took account of the direction of differences in levels of recognition for descriptions but not magnitude. For this reason they are less sensitive to patterns in the data than statistics used for the overall analyses. For example it is likely that they inflate the number of 'neutral' classifications by not taking into account the magnitude of any differences. Memory for the language of the description was not classified in this way for three reasons. Firstly, only three subjects reported using a verbal strategy. Secondly, scoring memory for the language of the descriptions without taking into account the magnitude of any differences in recognition proved to be very difficult. Thirdly, the spatial mental model account predicts high verbatim recognition scores (though only for indeterminate descriptions) resulting in a considerable overlap between the two classifications.

Memory for a spatial mental model was indicated by an interaction between determinacy and type of recognition; gist recognition should be higher than verbatim for determinate descriptions and verbatim higher than gist recognition for indeterminate descriptions (Mani & Johnson-Laird, 1982). For purposes of scoring this was considered as two criteria. If a person's performance was consistent with at least one of these criteria (e.g. determinate gist was higher than verbatim recognition) this was scored as memory 'for' a spatial mental model. If performance was inconsistent with at least one of these criteria (e.g. determinate verbatim was higher than determinate gist) it was scored as 'against' memory for a spatial mental model. If neither criteria was supported by recognition performance, or if one criterion was consistent and the other inconsistent with recognition performance, this was scored as 'neutral' with respect to memory for a spatial mental model.

Determining memory for an episodic construction trace (i.e. for the construction of a spatial mental model) was based on whether reordering a description reduced recognition performance (Payne, 1993). It was scored 'for' if verbatim recognition was higher for stable descriptions than for reordered descriptions. It was scored 'against' if reordered verbatim was higher than stable verbatim recognition. If stable and reordered verbatim recognition were equal this was scored as 'neutral'.

Figure 4.12 shows the classification of the subjects according to whether they showed evidence of remembering a spatial mental model broken down by primary strategy described in the retrospective protocol. The data are consistent with the view that the majority of subjects exhibited some memory for a spatial mental model. There is little evidence of an interaction between the primary strategy reported and memory for a spatial mental model (except possibly for those reported strategies classified as verbal). One surprising finding is that three subjects showing patterns of recall going against memory for a spatial mental model reported visual imagery for individual objects.

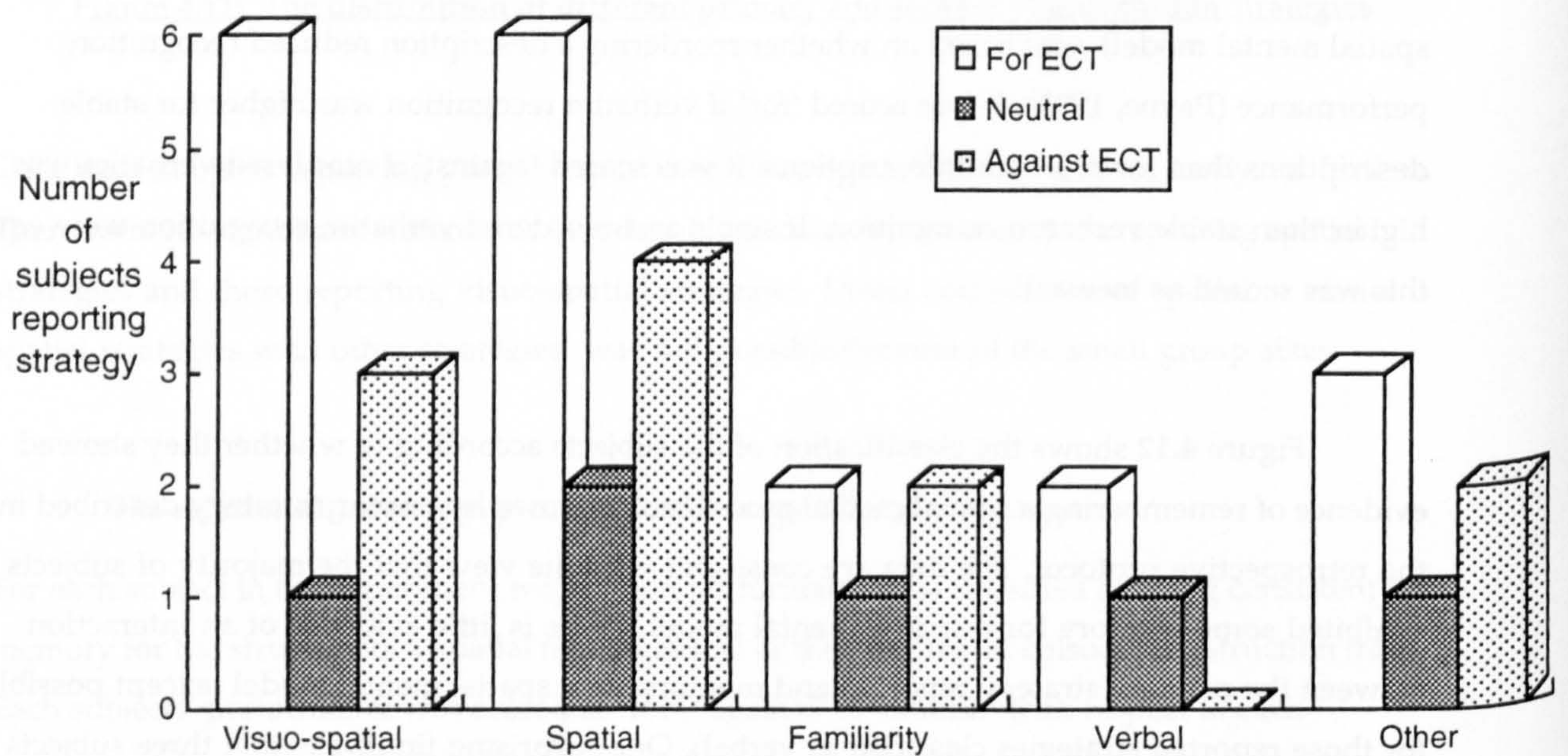




**Figure 4.12: Evidence of memory for a spatial mental model (SMM) by primary strategy.**

A possible explanation is that they had memorable or vivid visual imagery, but did not image the spatial or structural relationships between items in the description. This kind of occurrence is quite frequent when reading for pleasure. People often spontaneously experience vivid visual imagery, but it is relatively rare for that imagery to accurately reflect the spatial relationship contained in the story they are reading.

Figure 4.13 shows evidence of memory for an episodic construction trace according to the primary strategy reported by a subject.



**Figure 4.13: Evidence of memory for an episodic construction trace (ECT) by primary strategy.**



As might be expected memory for the episodic construction trace shows a similar pattern to memory for a spatial mental model, except that there are fewer 'neutral' classifications. One important difference is that two of the three people reporting verbal strategies show evidence of basing recognition on memory for an episodic construction trace.

Looking at the recognition performance in detail supports the pattern of results obtained in the overall analysis. It also appears to support the view that subjects who report visuo-spatial or spatial strategies are likely to have constructed and remembered spatial mental models, or an episodic construction trace (or both). However, it is difficult to reach any strong conclusions where classifications of recognition test performance and retrospective protocols result in such small cell sizes.

### *Working memory and recognition performance*

A sample of subjects were also tested for their backward digit span and reading span in order to determine if either of these measures of working memory capacity was a good predictor of recall (Daneman & Carpenter, 1980). Thirty-five subjects were tested for backward digit span. Twelve subjects were also tested for reading span. Backward digit span and reading span appeared to be independent ( $Rho = 0.039$ ). Backward digit span correlated significantly with gist recognition ( $Rho = 0.398$ , two-tailed  $p = 0.02$ ) but not verbatim recognition ( $Rho = 0.044$ , two-tailed  $p = 0.80$ ). This is inconsistent with a strategy where subjects engage in verbal rehearsal of the sentences after reading them. It suggests, instead, that verbal working memory resources (e.g. the articulatory loop) may be involved in holding verbal information while a spatial mental model is constructed. This is consistent with several other findings in the mental models literature (Baddeley & Wilson, 1988; de Vooght & Vandierendonck, 1993; Morra, 1989; Oakhill & Johnson-Laird, 1984).

Reading span showed a negative relationship with scores for determinate descriptions on the phase 1 matching task and with determinate gist recognition. This relationship was significant between reading span and determinate matches ( $Rho = -0.775$ ,  $N = 12$ , two-tailed  $p = 0.02$ ). For reading span and determinate gist recognition this trend did not reach significance ( $Rho = -0.524$ ,  $N = 12$ , two-tailed  $p = 0.12$ ). Reading span showed no evidence of a positive association with any of the remaining recognition scores. There is therefore no evidence that reading span is involved in the retention of spatial descriptions. In fact, the evidence suggests that subjects with high reading spans were worse at matching determinate descriptions in phase 1. One possible explanation is that subjects with low readings spans compensated for this by concentrating harder or taking longer to read the descriptions. Subjects with high reading spans may have felt matching the determinate descriptions to be sufficiently easy not to expend much effort (remember that all the subjects scored over fifty percent in the phase 1 diagram matching task).



### 4.1.3 Discussion

When people are asked to compare spatial descriptions with diagrams they learn information about those spatial descriptions. If, at a later stage, they are given a surprise recognition test it has been proposed that they may show evidence of three types of memory; memory for the language of the description, memory for the construction of a spatial mental model or memory for a spatial mental model itself. In this experiment there was an interaction between type of recognition memory and the determinacy of the description. For determinate descriptions gist recognition was higher than verbatim recognition. For indeterminate descriptions verbatim recognition was higher than gist.

This finding replicates the crossover interaction obtained by Mani and Johnson-Laird (1982). It partially contradicts the results obtained by Payne (1993) even though the materials were closely based on those used in his studies. Mani and Johnson-Laird explained their findings as supporting the view that people attempt to construct a spatial mental model but abandon this attempt if becomes too difficult (e.g. when they realize the description is indeterminate). Hence for determinate descriptions recognition is based primarily on memory for a spatial mental model which can not differentiate between the original and inferable description. When people realize that a description is indeterminate and abandon spatial mental model construction Johnson-Laird argues that they concentrate on remembering the language of the description instead (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). However, the high level of gist recognition for the reordered indeterminate descriptions in the present experiment suggests that some people are able to construct a spatial mental model of an indeterminate description. Johnson-Laird (1983) discusses a number of ways a spatial mental model can represent indeterminacy. An alternative explanation of verbatim recognition scores for indeterminate descriptions being higher gist recognition scores also suggests itself. People may continue to attempt to construct spatial mental models, but, in doing so, they take longer and have to reread the original description several more times. Thus raised verbatim recognition for indeterminate descriptions may be the result of a specific attempt to learn the description or a by-product of the difficulty of spatial mental model construction for indeterminate descriptions. While it is likely that both accounts have something to offer, the relatively high gist recognition scores for the indeterminate descriptions favour the latter.

Reordering the original descriptions reduced verbatim recognition levels, however, this effect was significant only for the matching descriptions. Examination of performance for the non-matching condition suggests that these descriptions may have varied greatly in difficulty. If this was the case any differences between these descriptions would be harder to detect. When the original description is reordered the overlap between the episodic construction trace of the



original description and the target in the recognition test is reduced (Payne, 1993). So the evidence from this experiment shows only mixed support for the view that people remember an episodic construction trace. Analysis of individual differences shows that many people do show evidence of remembering an episodic construction trace and that there may be some association between memory for a spatial mental model and memory for an episodic construction trace. One other finding also supports the view that some people remember constructing a spatial mental model. Gist and verbatim recognition scores are correlated with each other. Memory for a spatial mental model does not support verbatim recognition and verbatim recognition is usually higher than gist recognition. It is therefore likely that at least some gist recognition can be accounted for by memory for an episodic construction trace (the stable original is ranked first and then it is realized that the inferable describes the same configuration and it is ranked second).

The argument that people remember an episodic construction trace depends to some extent on a lack of evidence that people remember the language of the descriptions. Several findings indicate support for the assertion that people are not basing recognition on a representation of the language of the text. Firstly, gist recognition for determinate descriptions exceeds verbatim recognition. Secondly, the observation that reordering depresses verbatim recognition for any of the descriptions is difficult to account for in terms of memory for the sentences of the description. Hundreds of verbal learning experiments have been carried out in psychology which show that people can and do recognize individual sentences taken from a text. Lastly, only three subjects reported trying to remember the language of the descriptions in any way (and two of those subjects showed evidence of remembering an episodic construction trace).

The overall pattern of results provides some support for the hypothesis that people remember both the construction of a spatial mental model and the structure of the spatial mental model. The lack of differences between the one diagram and three diagram condition makes discussion of the influence of repeated presentation difficult. It appears that presenting the spatial descriptions more than once does not greatly influence quantity or quality of learning. The relationship between reading span and performance may shed some light on this observation. Reading span correlates negatively with determinate performance on the matching task. It is possible that those with higher reading spans were more confident on the matching task and hence spent less time reading the easier, determinate descriptions. If this is the case it suggests that in self timed tasks such as these people read the descriptions until they feel comfortable that they can remember them. Presenting descriptions three times does not significantly help them because they spend less time on each description than those in the single presentation condition.

The positive relationship between gist recognition and backward digit span may seem surprising. The lack of a relationship between verbatim memory and backward digit span also



suggests that few if any of the subjects were actively attempting to learn the language of the descriptions. However, a number of other studies have shown a relationship between measures of short-term verbal memory and memory for a spatial mental model (Baddeley & Wilson, 1988; de Vooght & Vandierendonck, 1993; Morra, 1989; Oakhill & Johnson-Laird, 1984). As high digit span appears to improve gist retention, this suggests that the articulatory loop may help hold information from the description and so 'free up' general working memory resources for spatial mental model construction and consultation (also see section 2.2.2).

This experiment provides support for the notion that people can and do remember the structure of a spatial mental model (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). It also provides partial support for the hypothesis that people remember the process of construction of a spatial mental model. Unfortunately there is also evidence that memory for descriptions may be strongly influenced by the difficulty of the particular descriptions used in this and other experiments (Mani & Johnson-Laird, 1982; Payne, 1993). The next two experiments (5 and 6a) examine memory for individual descriptions in more detail.



## 4.2 Experiment 5

This experiment investigates free recall in a drawing task for the eight spatial descriptions used in Experiment 4. Immediately after taking part in Experiment 4 thirty people were asked to draw out the items from each spatial description in positions consistent with the original description they had read. The experiment had three main aims. Firstly, to discover whether determinacy influences recall as well as the recognition of spatial descriptions. Secondly, to examine the relationship between spatial mental model construction order and drawing order using simpler stimuli than those used in Chapter 3. Thirdly, to investigate the relationship between spatial mental model construction order and levels of recall.

Experiment 4 has provided strong evidence that many, if not all, of the subjects constructed or attempted to construct a spatial mental model consistent with each of the spatial descriptions they read. The account of spatial mental model construction outlined in Chapter 2 predicts that both drawing order and probability of recall should reflect the order in which items are integrated into a spatial mental model (Payne, 1993; Taylor & Tversky, 1992a; Tversky, 1991). These predictions do not differ for accounts of recall based on the structure of a spatial mental model and those based on memory for its construction e.g. in the form of an episodic construction trace. The classic account of memory for a spatial mental model (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982) might also predict a difference in drawing order and recall between determinate and indeterminate descriptions. Recall for determinate descriptions should reflect the construction and retention of a spatial mental model. Recall for indeterminate descriptions may show the same pattern or it may reflect memory for the language of the description.

A lack of a drawing order effect, or lack of a relationship between spatial mental model construction order and probability of recall would call into question the interpretation of the studies reported in Chapter 3. It would also support interpretations of drawing recall that postulate a representation such as an image that does not retain the order of its construction (Kulhavy et al., 1993; Kulhavy et al., 1993; Tversky, 1991). It would also be possible to obtain other effects related to the serial position of items in recall or drawing order. Obtaining a classic serial position effect (depressed recall or delayed drawing order for items in the middle of the description) or a recency effect might favour explanations of recall in terms of a propositional representation of the language of the descriptions. For a more detailed discussion of these arguments refer to Chapter 3.



### 4.2.1 Method

#### Subjects

Thirty of the thirty-six subjects from Experiment 4 also agreed to take part in this experiment.

#### Materials

For each of the eight spatial descriptions used in Experiment 4 an empty grid made up of six blank boxes was prepared. The six boxes were made up of either two rows and three columns or three rows and two columns as appropriate for the permissible spatial layouts of each description. Each grid was given one of the headings used in the eight original descriptions (e.g. 'Tableware' or 'Transport') indicating the appropriate items to be entered into the grid.

#### Procedure

This experiment was carried out as soon as each subject had completed Experiment 4. None of the subjects in Experiment 4 were expecting to be tested for a second time. The delay between presentation and test was thus determined by how long each subject took to complete Experiment 4. Subjects took approximately thirty to forty minutes to complete phases 2 and 3 of Experiment 4. Where subjects were tested for backward digit span or reading span these tests were performed after the Experiment 5 had been completed. Subjects were given a series of empty three by two grids labelled for each of the eight description categories from Experiment 4. They were asked to produce spatial configurations of objects consistent with the original descriptions received in Experiment 4. The majority of subjects chose to write the names of items into the grids, though several subjects preferred to sketch the objects. The experimenter determined in which order the drawings were made by prompting each subject with the category type of each description in a random order. Within each configuration a record of the drawing order of the objects was made by the experimenter (refer to Experiment 1 for the procedure used to record drawing orders).

### 4.2.2 Results

#### Drawing recall of spatial descriptions

##### *Scoring*

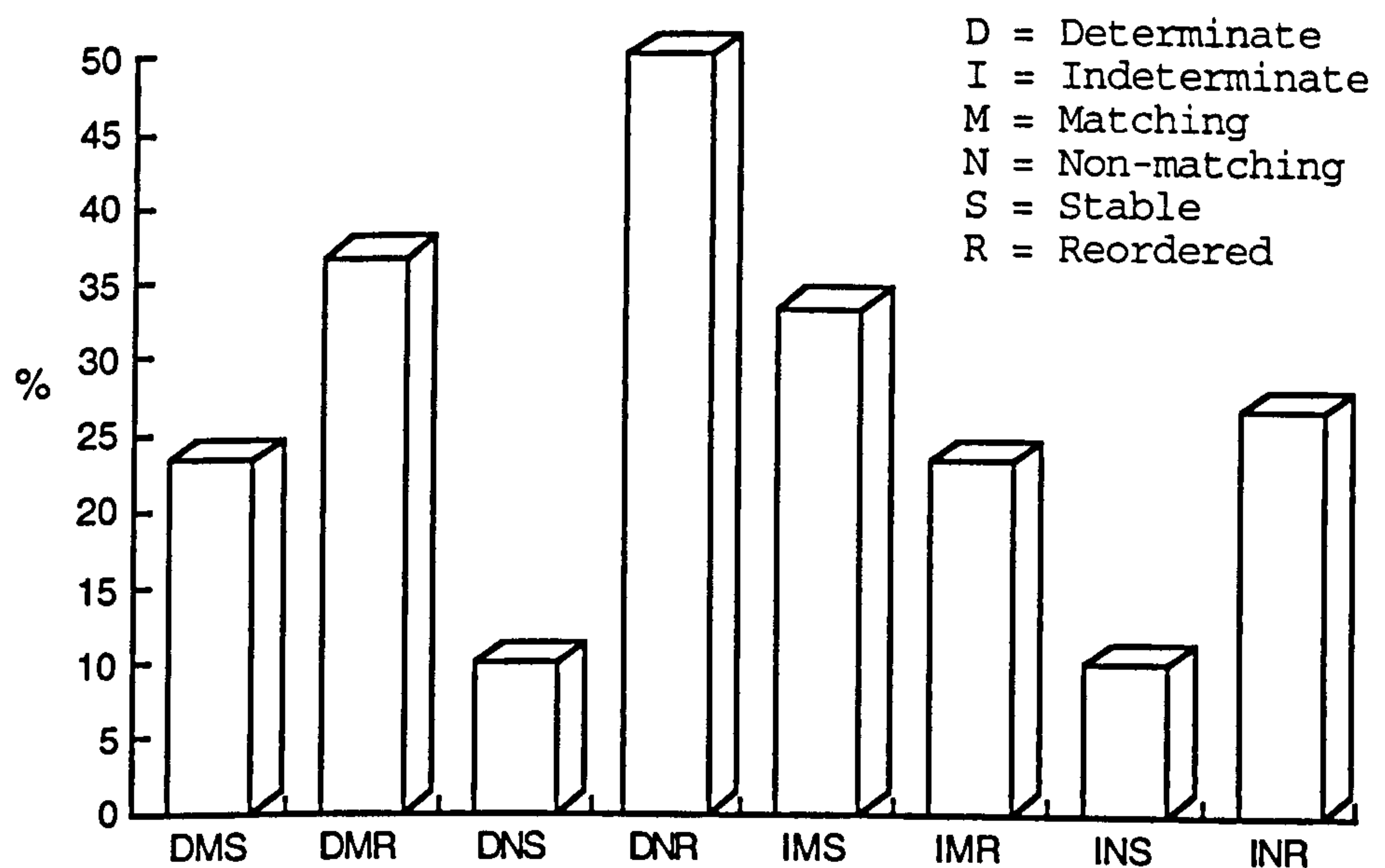
Subjects' drawings were scored on a number of measures; object recall, relation recall and drawing order. Object recall was scored if a given object was present in the appropriate description. Relation recall was scored if, for a given sentence in the original description, both objects were located correctly with respect to each other. For each description a record was made if all objects



were correctly placed in the drawing (a completely correct drawing received a score of '1', an incomplete or inaccurate drawing received a score of '0'). In addition, for each description drawing order was measured by calculating Kendall's tau for each drawing order with respect to the predicted order of spatial mental model construction. For an illustration of how spatial mental model construction order was determined refer to Chapter 2 (note the spatial descriptions contained only asymmetric spatial relations). In the case of these descriptions the predicted spatial mental model construction order is '2-1-3-4-5' where each number indicates the serial position of that item in the original description.

### *The influence of differences between descriptions on recall*

During the analysis of Experiment 4 it became apparent that there were strong effects of item difficulty on the overall results, in particular for the non-matching descriptions. Many subjects reported that one description was much easier to 'visualize' than the others ('Landmarks'), whilst other descriptions were consistently reported as being more difficult to imagine ('Garden Tools' and 'Plants' in particular). These impressions are supported by levels of recall on the drawing task. Figure 4.14 below shows the percentage of completely correct drawings by condition (note that each condition is associated with only a single description and hence category). Of particular interest is the difference between the reordered and stable descriptions despite the fact that this variable is not implicated in the drawing task.



**Figure 4.14: The percentage of correct drawings by description type.**

These scores were compared across items using Wilcoxon tests for all thirty subjects. The proportion of correct drawings for the 'Garden Tools' (DNS), 'Landmarks' (DNR) and 'Plants' (INS) descriptions were significantly different from the overall mean ( $Z = -3.03$ , two-tailed  $p = 0.0025$ ;  $Z = -2.51$ , two-tailed  $p = 0.012$  and  $Z = -3.75$ , two-tailed  $p = 0.0002$  respectively). These results support the interpretation of the effect of reordering in Experiment 4. It may be that the

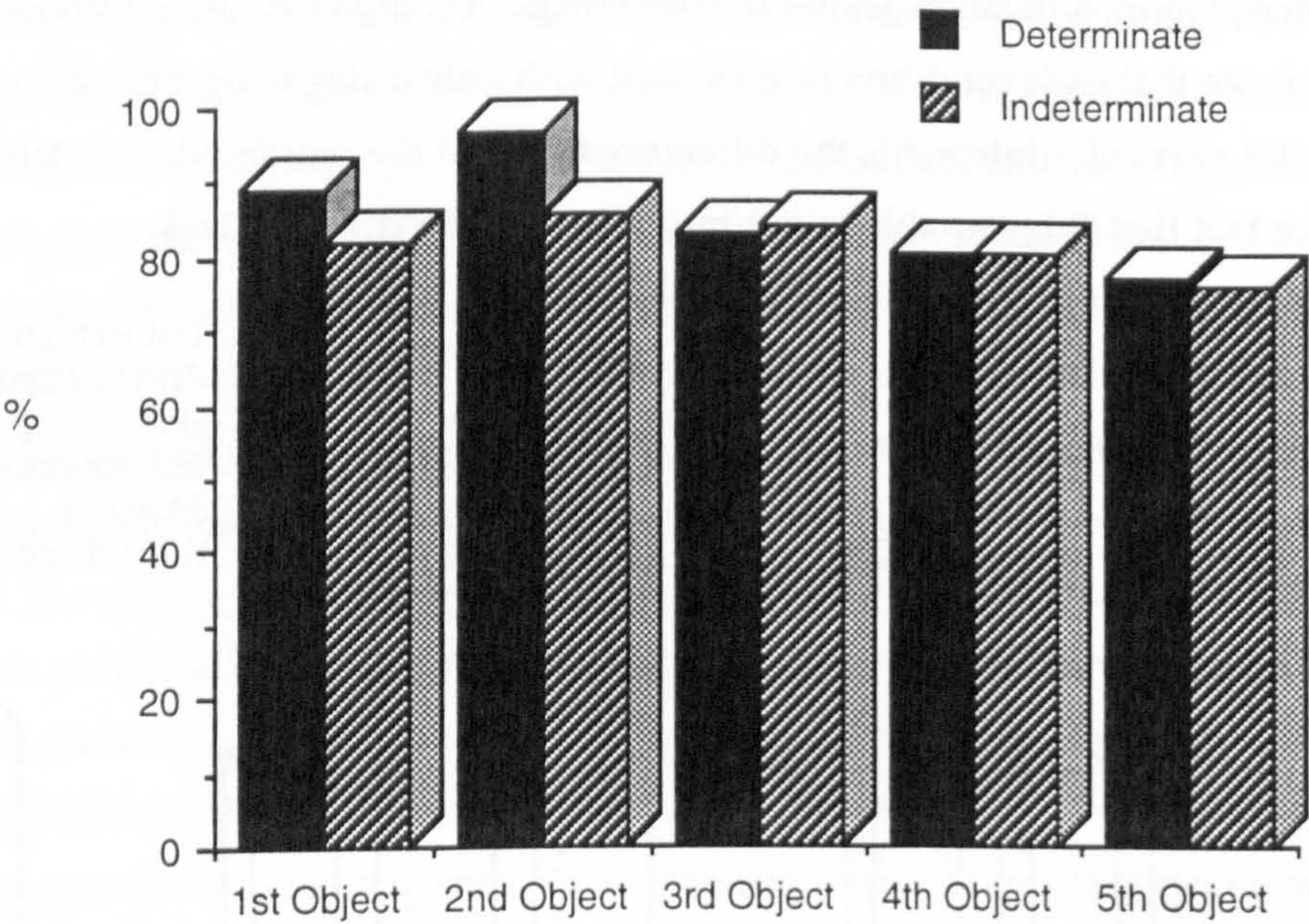


lack of an effect of reordering for the matching descriptions may be caused by differences in memorability between the descriptions. This possibility is explored more fully in Experiment 6a.

For the majority of the remaining analyses data are averaged over several descriptions to minimize the differences due to the content of the descriptions themselves. However, the remaining analyses involve comparisons within descriptions for the order or probability of recall. For this reason the large differences in recall between particular descriptions should not greatly influence the results.

*Overall object recall*

Percentage object recall for determinate and indeterminate descriptions was calculated for each object in the eight descriptions. From these a mean level of recall for the item in each serial position was obtained. Averaging recall over all eight descriptions reduces the influences of distinctiveness of individual items or association between pairs of items. These are summarized by item in Figure 4.15.



**Figure 4.15: Percentage recall of objects by order and description type.**

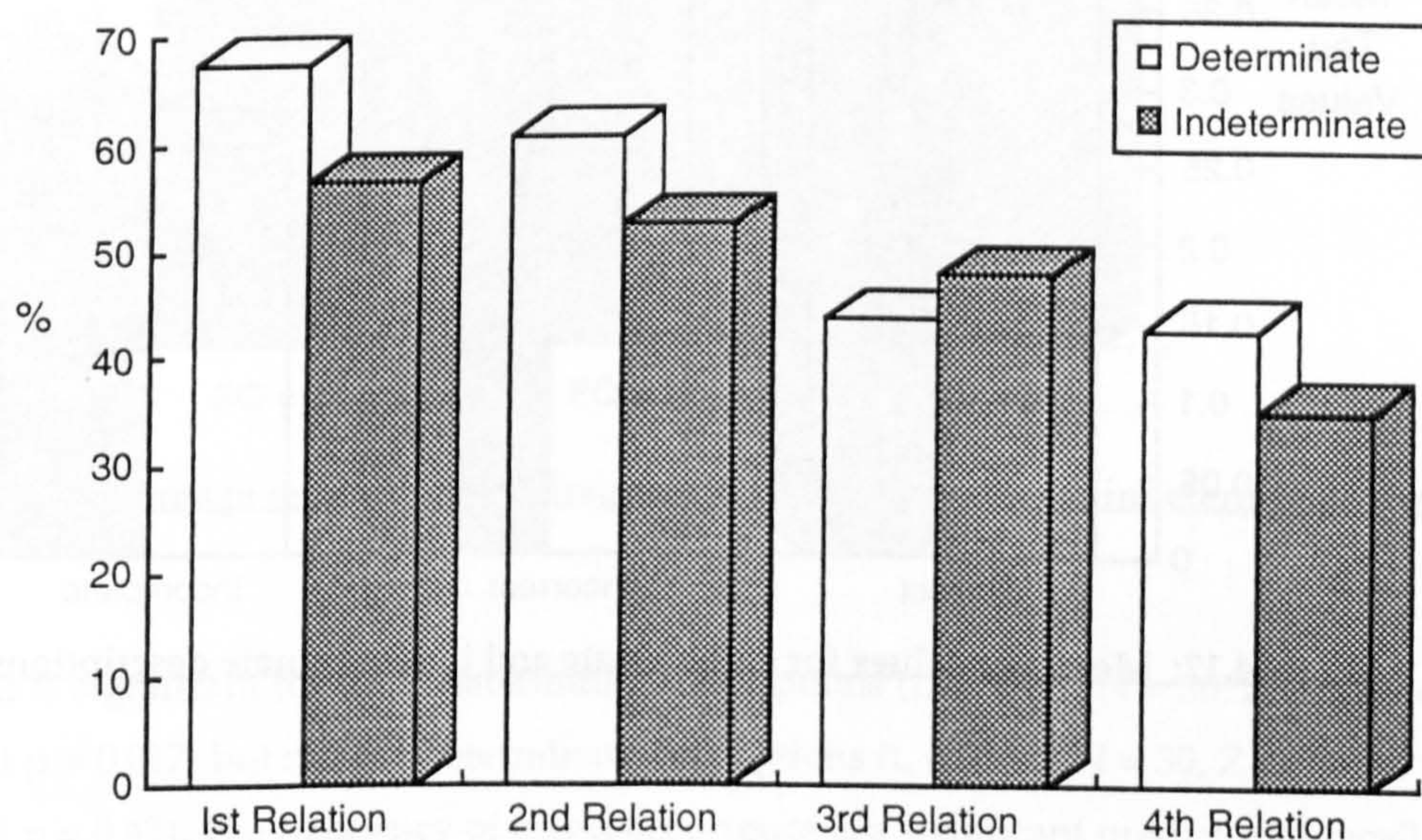
The predicted pattern of results for the mental models and episodic construction trace account (mean recall decreasing in the order second object, first object, third object, fourth object and fifth object) was analyzed using the Page test for ordered alternatives (Siegel & Castellan, 1988). This was found to be significant for the overall pattern ( $L = 1450$ ,  $N = 30$ ,  $Z$  estimate = 3.65, one-tailed  $p < 0.0002$ ) as well as separately for the determinate ( $L = 1437.5$ ,  $N = 30$ ,  $Z$  estimate = 3.20, one-tailed  $p < 0.0013$ ). There was also a trend in this direction for the indeterminate descriptions, though this did not reach significance ( $L = 1393.5$ ,  $N = 30$ ,  $Z$  estimate = 1.59, one-tailed  $p = 0.057$ ).



The findings for object recall thus seem to support the prediction that the probability of recalling an object from early in the spatial mental model construction is higher than that for objects later in the construction process. This relationship is particularly marked for the first two objects in a determinate description. This supports the view that later objects added to a spatial mental model improve recall for objects, because spatial mental model construction for the determinate descriptions proceeds until all five objects are added to the model. Spatial mental model construction for the indeterminate descriptions is likely to break down when the indeterminacy is reached (at the third, fourth or fifth object). This may also explain the weaker relationship between spatial mental model construction order and recall for the indeterminate description.

### *Overall relation recall*

The percentage recall of each relation for determinate and indeterminate descriptions was calculated for all descriptions. These are summarized in Figure 4.16 below.



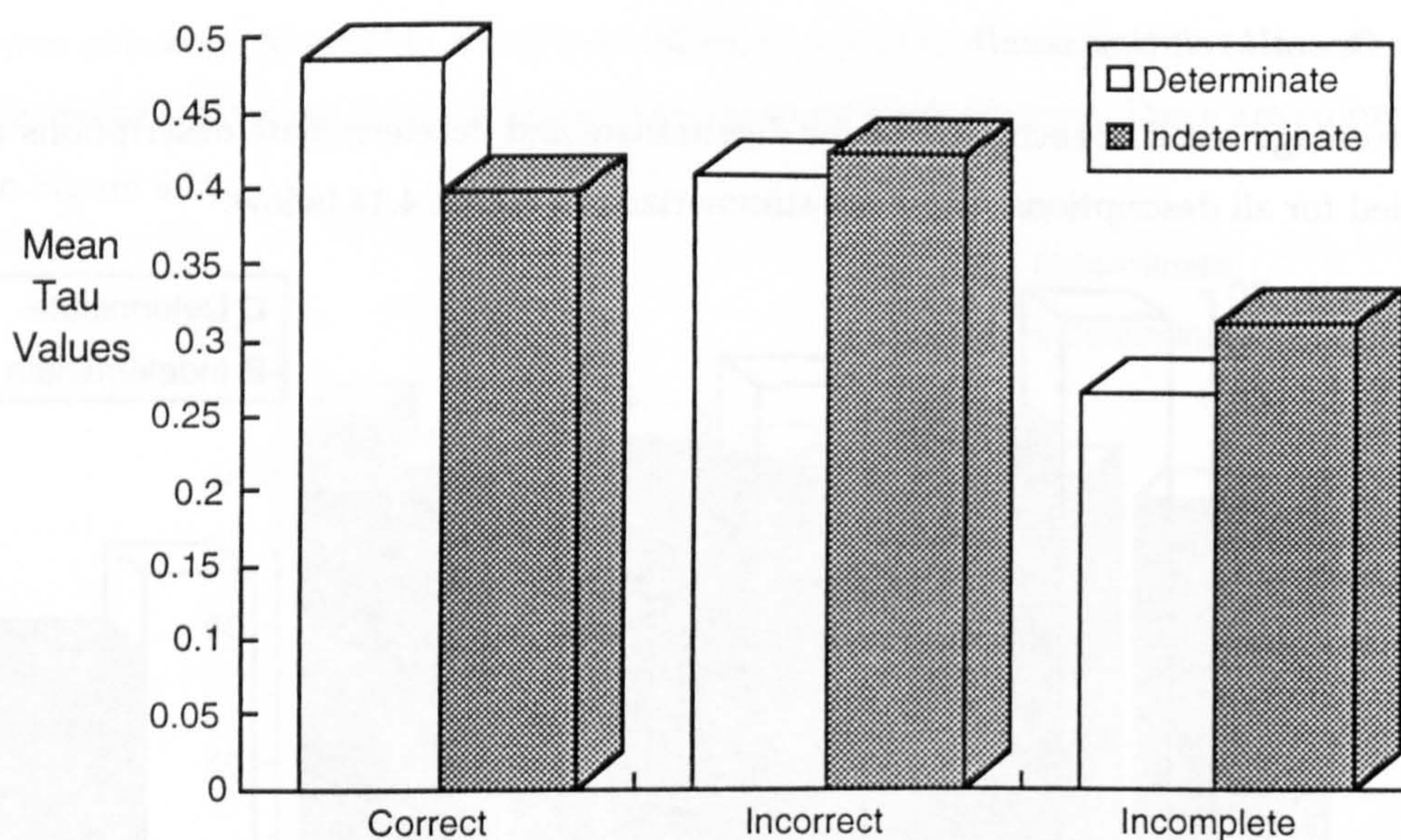
**Figure 4.16: Percentage recall of spatial relations by order and description type.**

These data was analyzed using the Page test for ordered alternatives. The predicted trend is for recall to decrease steadily from the first to last spatial relation. The first relation is predicted to be recalled best because it is between the first and second objects integrated into the spatial mental model. The resulting test statistic was significant ( $L = 829$ ,  $N = 30$ ,  $Z$  estimate = 5.00, one-tailed  $p < 0.0001$ ). When analyzed separately for the two description types the trend is significant for both determinate ( $L = 822.5$ ,  $N = 30$ ,  $Z$  estimate = 4.59, one-tailed  $p < 0.0001$ ) and for indeterminate descriptions ( $L = 793$ ,  $N = 30$ ,  $Z$  estimate = 2.72, one-tailed  $p = 0.033$ ). Again the results are in favour of the predicted relationship between spatial mental model construction and recall. This suggests that not only are early objects more likely to be recalled, but that the spatial relationships between them are better remembered.



### Overall drawing order

It was possible to calculate the Kendall's tau between drawing order and spatial mental model construction order for drawings where subjects had recalled at least four objects from the original description. The tau values for the incomplete drawings provide a generous estimate of the association possible by chance. If tau values for complete drawings (whether correct or incorrect) exceed those of incomplete drawings it is possible to argue beyond reasonable doubt that they would also exceed those obtainable by chance or by the intrusion of preferred drawing orders. Figure 4.17 shows the mean tau values for determinate and indeterminate descriptions according to whether subjects drew the description correctly, incorrectly or incompletely (i.e. only four items were recalled).



**Figure 4.17: Mean tau values for determinate and indeterminate descriptions.**

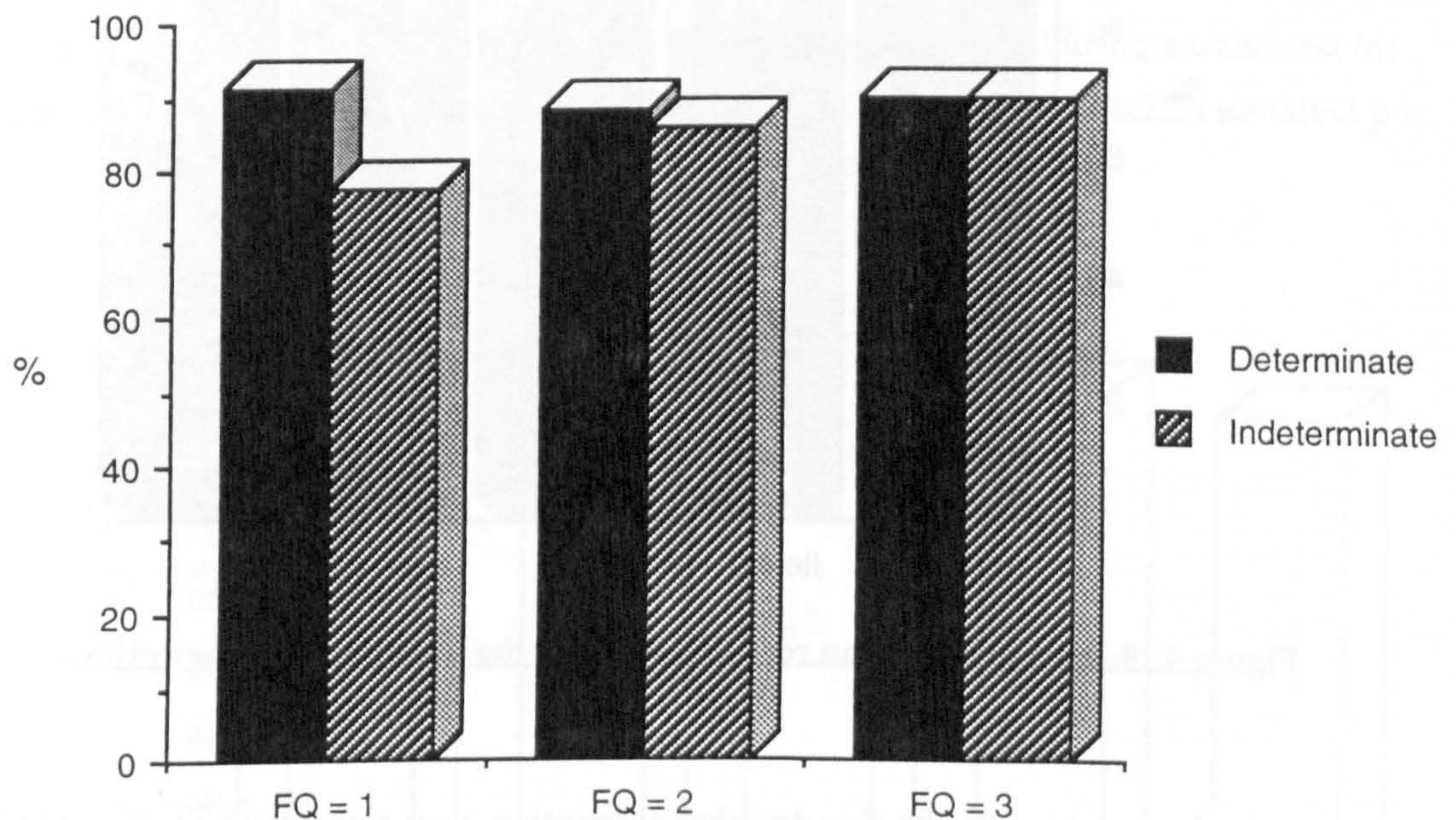
Differences between these tau values were examined using related t-tests. Over all eight descriptions the tau values were higher for the correct and incorrect descriptions where all five objects were recalled than when only four objects were recalled ( $t = 2.52$ , d.f. = 20, one-tailed  $p = 0.0093$ ). Tau values for determinate and indeterminate descriptions do not differ significantly. These tau values are far higher than would be expected for any chance relationship between spatial mental model construction order and drawing order. This suggests that drawing orders are influenced by the order of spatial mental model construction.

### Frequency of object occurrence

A serious methodological problem arises when looking at overall levels of recall (where mean object recall is calculated over all eight descriptions for each subject). Predicted spatial mental model construction order is partially confounded with frequency (the number of times an object



occurs in the description). Within each description one object is mentioned three times and another object is mentioned twice. In general, early objects are mentioned more often than late objects. For this reason frequency differences are unlikely to be wholly responsible for the drawing order effects obtained, though they may contribute to them. It should also be noted that the fifth object in the description never occurs more than once yet is most often drawn last and is recalled worst. Figure 4.18 shows the overall relationship between frequency and probability of recall for determinate and indeterminate descriptions. The figures for objects mentioned only once (FQ = 1) are the earliest mentioned object mentioned once in each description.



**Figure 4.18: Percentage recall by object frequency for determinate and indeterminate descriptions.**

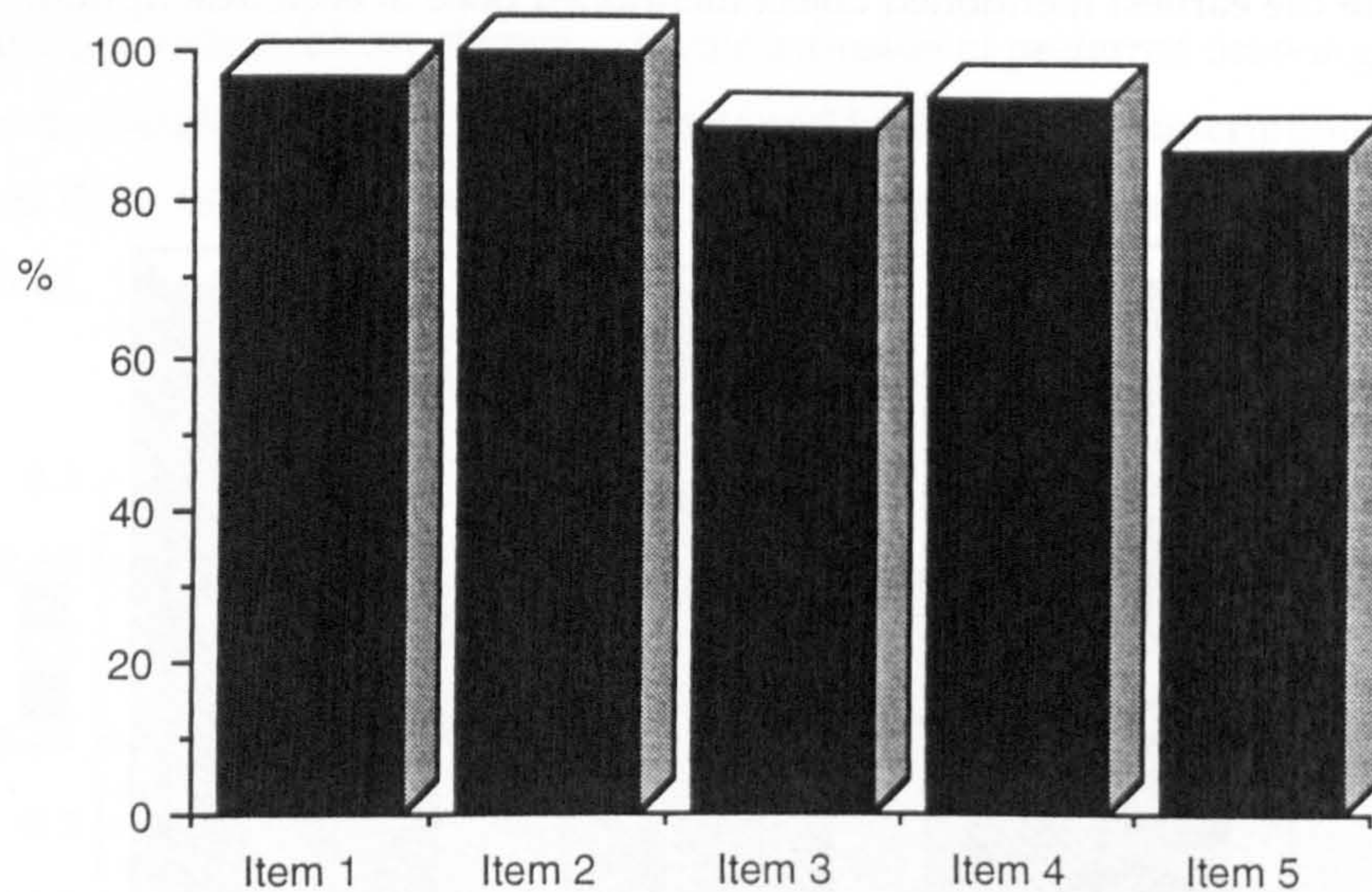
This trend is significant for the indeterminate descriptions ( $L = 376.5$ ,  $N = 30$ ,  $Z$  estimate = 2.13, one-tailed  $p = 0.017$ ) but not the determinate descriptions ( $L = 358.5$ ,  $N = 30$ ,  $Z$  estimate = -0.19, one-tailed  $p = 0.42$ ). The frequency of object occurrence is a significant predictor of recall for indeterminate descriptions, but not determinate descriptions. This suggests that the recall of determinate descriptions is not significantly contaminated by the number times each object is mentioned in the original description.

For individual descriptions the spatial mental model construction order and the order predicted by frequency can differ. The remaining results sections examine drawing order and probability of recall for individual descriptions where predictions from frequency and spatial mental model construction order differ. The best recalled determinate and indeterminate description are used to illustrate these patterns.



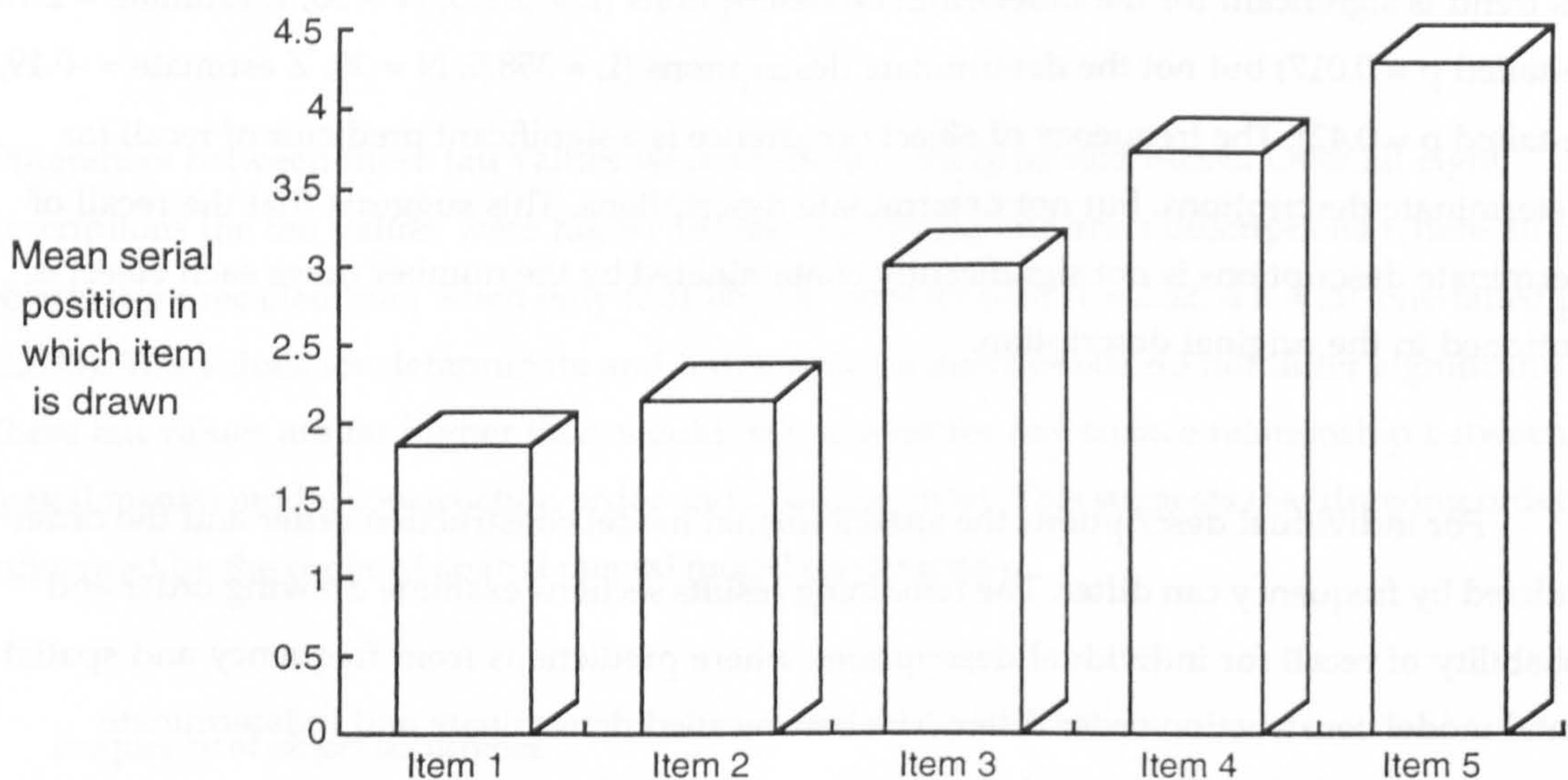
### *Memory for the determinate 'landmarks' description*

This was the single best remembered description in the drawing task. Figure 4.19 shows mean percentage recall by item for the 'Landmarks' description. The trend depicted in the graph differs from that predicted by frequency (the most frequently mentioned landmarks were item 1 followed by item 4). This trend was analyzed using a trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). Even though for many subjects recall was at ceiling this trend was still significant ( $N = 30$ ,  $Z$  estimate = 2.18, one-tailed  $p = 0.015$ ).



**Figure 4.19: Percentage mean recall by item for the 'Landmarks' description.**

The drawing orders of items in the 'Landmarks' description were also analyzed. This analysis was carried out between items using the Jonckheere trend test. Figure 4.20 shows the mean drawing order for all five items in the description. The analysis excludes those subjects who drew less than four items.



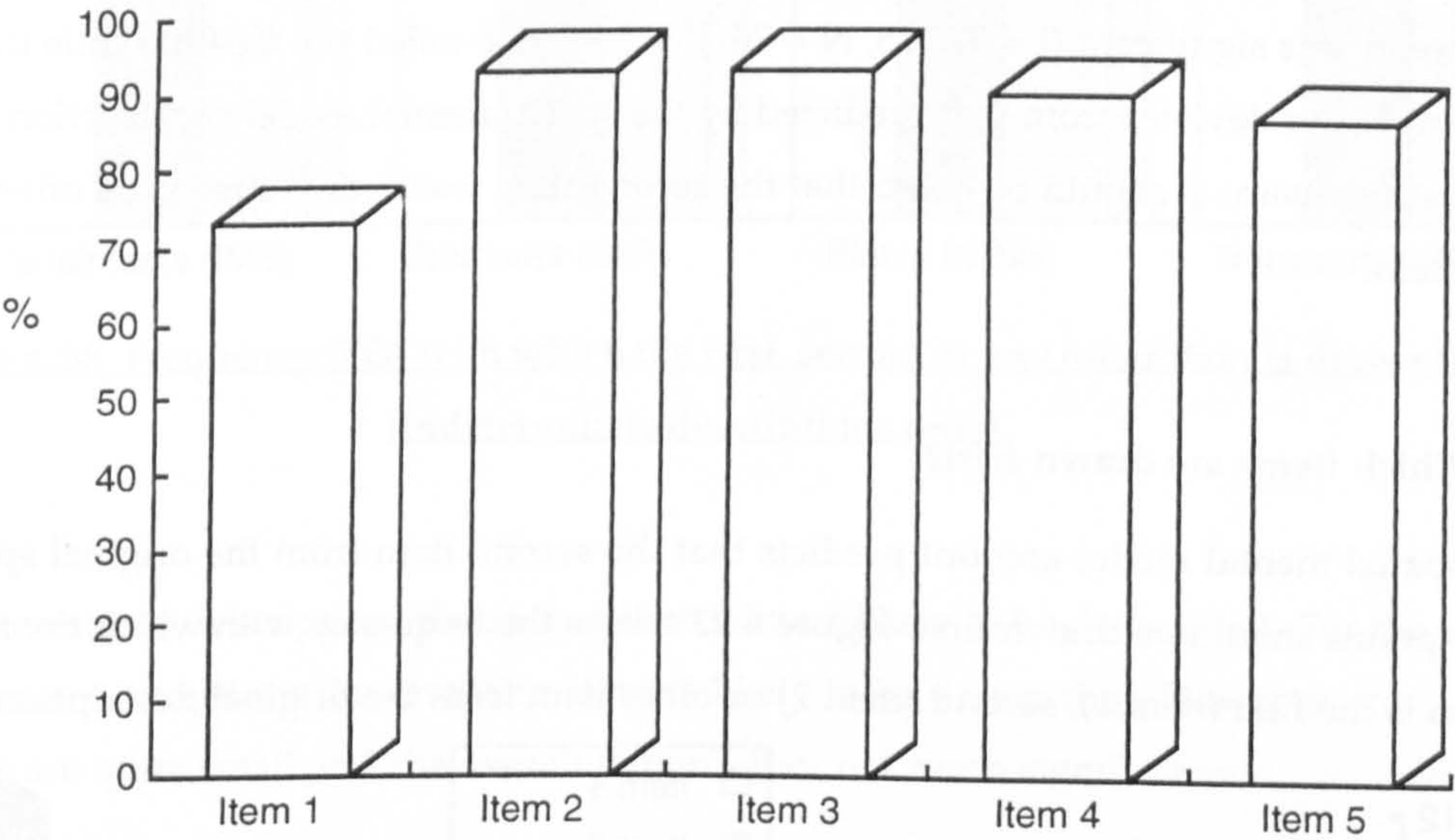
**Figure 4.20: Mean serial position in which items in the 'Landmarks' description were drawn.**



This trend was significant ( $J = 5685.5$ ,  $N = 27$ ,  $J^* = 7.93$ , one-tailed  $p < 0.0001$ ). While mean drawing order deviates from that predicted by the spatial mental model construction account for the first two items it should be noted that the second item was drawn first more often than the first item.

*Memory for the indeterminate 'condiments' description*

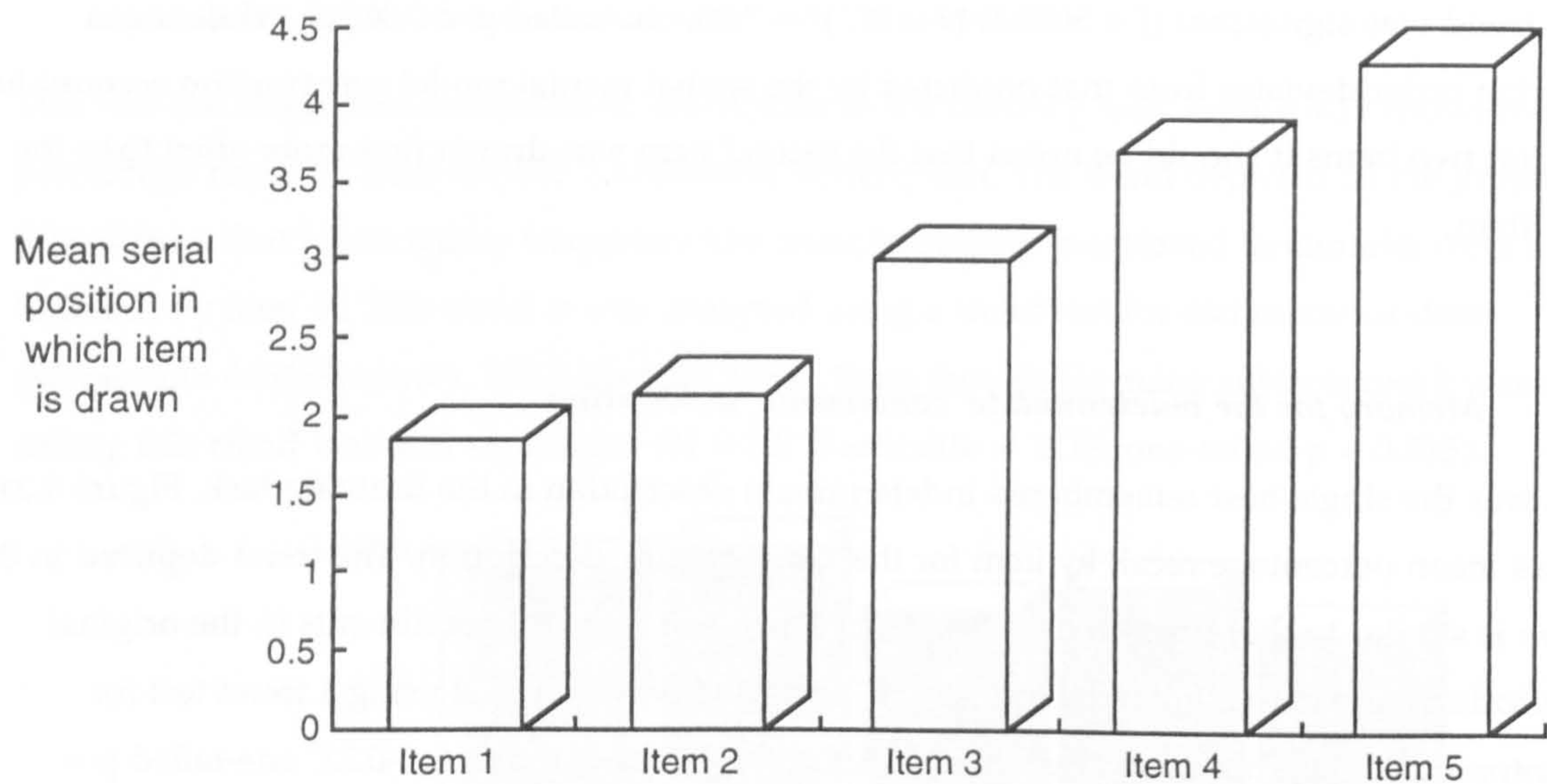
This was the single best remembered indeterminate description in the drawing task. Figure 4.21 shows mean percentage recall by item for the 'Condiments' description. The trend depicted in the graph is similar to that predicted by frequency (the most frequent condiments in the original description were item 2 followed by item 3). This trend was analyzed using a trend test for dichotomous data. This trend was not significant ( $N = 30$ ,  $Z$  estimate =  $-0.22$ , one-tailed  $p = 0.41$ ).



**Figure 4.21: Percentage mean recall by item for the 'Condiments' description.**

The drawing orders of items in the 'Condiments' description were also analyzed. This analysis was carried out between items using the Jonckheere trend test. Figure 4.22 shows the mean drawing order for all five items in the description. The analysis excludes those subjects who drew less than four items.



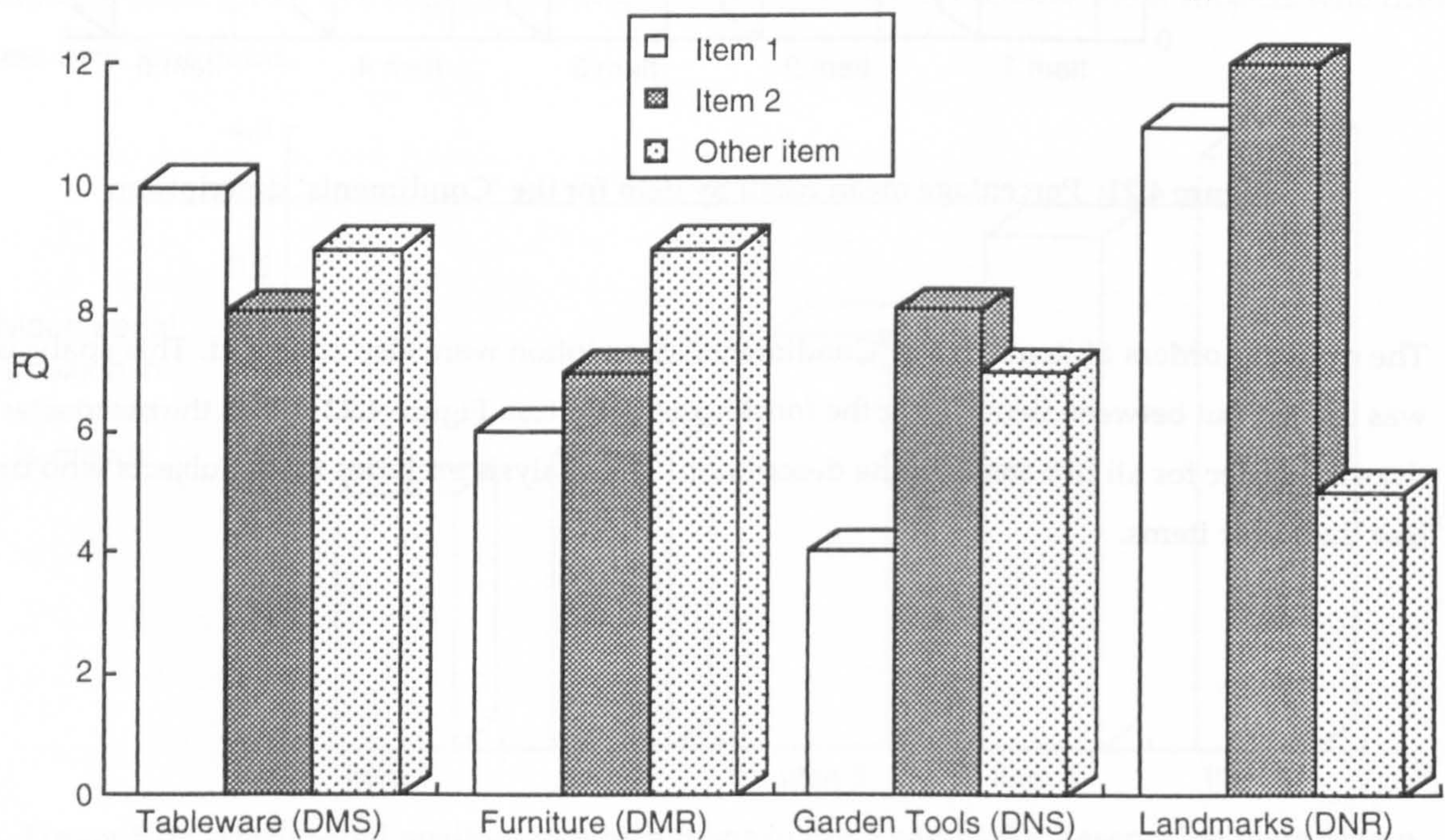


**Figure 4.22: Mean serial position in which items in the 'Condiments' description were drawn.**

This trend was significant ( $J = 3783.5$ ,  $N = 26$ ,  $J^* = 1.66$ , one-tailed  $p = 0.049$ ). While mean drawing order deviates from that predicted by the spatial mental model construction account for the first two items it should be noted that the second item was drawn first more often than the first item.

### Which items are drawn first?

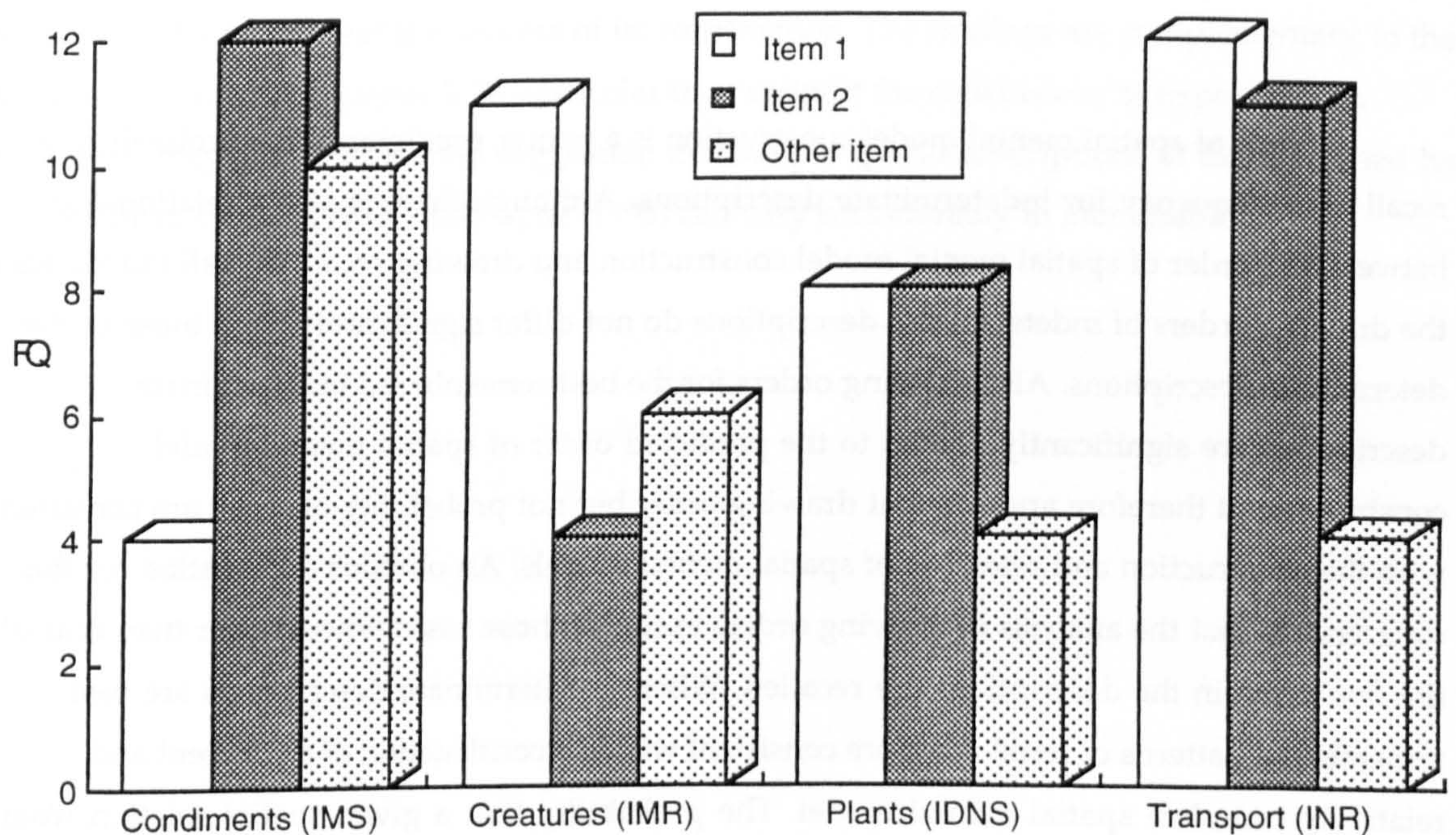
The spatial mental model account predicts that the second item from the original spatial descriptions should be drawn first. Figure 4.23 shows the frequency with which the first item drawn is the first (item 1), second (item 2) or other item from the original description.



**Figure 4.23: Frequency (FQ) with which the first, second or any other item is drawn first (determinate descriptions only).**



Figure 4.24 shows the frequency with which items are drawn first for the indeterminate descriptions.



**Figure 4.24: Frequency (FQ) with which the first, second or any other item is drawn first (indeterminate descriptions only).**

For the determinate descriptions the second item is most likely to be drawn first. For the indeterminate descriptions the first item is more likely to be drawn first. However, the differences are quite small and the overall pattern does not reach significance.

### 4.2.3 Discussion

There is a strong relationship between drawing recall and the predicted order of spatial mental model construction for determinate descriptions. This is shown both in the drawing orders people adopt when asked to draw out the spatial descriptions they have read and probability that a particular spatial relationship or item is recalled. The earlier an item is integrated into a spatial mental model the more likely it is to be recalled, and, if recalled, the earlier it is drawn. This is consistent with research that proposes memory for determinate spatial descriptions is based on memory for a spatial mental model or for its construction (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982; Payne, 1993). It also supports other research that proposes that drawing a spatial description involves mentally reconstructing a spatial mental model of a learned environment (Taylor & Tversky, 1992a; Tversky, 1991). Two alternative explanations of ordinal patterns of recall can be rejected. There is no evidence of a recency effect or of depressed recall for



items in the middle of the determinate descriptions. Therefore an explanation in terms of a classic serial position effect, for example, for a list of items or for the propositions of a text is not supported ( Craik, 1970; Freebody & Anderson, 1986). Nor is the pattern of recall for *determinate* descriptions predicted by the frequency with which objects in a given description are mentioned. Therefore the results cannot be explained in terms of frequency.

Order of spatial mental model construction is a poorer predictor of the probability of recall than frequency for indeterminate descriptions. Although there is also a relationship between the order of spatial mental model construction and drawing recall, overall tau values for the drawing orders of indeterminate descriptions do not differ significantly from those of the determinate descriptions. Also drawing orders for the best remembered indeterminate descriptions are significantly related to the predicted order of spatial mental model construction. It therefore appears that drawing order but not probability of recall are consistent with the construction and retention of spatial mental models. An obvious explanation for this difference is that the analysis of drawing orders excludes those cases where fewer than four of the five items in the descriptions are recalled. Where indeterminate descriptions are best remembered patterns of recall are more consistent with the construction of a coherent and relatively complete spatial mental model. The probability that a given spatial relation from the original description is recalled is also related to the order of spatial mental model construction. This makes sense if the spatial properties of the drawing are influenced by memory for a spatial mental model. Where a complete or coherent spatial mental model was not constructed other factors, in particular frequency, determine whether an object is recalled. The influence of frequency on recall for indeterminate descriptions, but not determinate descriptions, suggests that people may be remembering the language of the description. Items mentioned in more than one sentence are more likely to be recalled. This is consistent with the interaction between determinacy and gist or verbatim recognition obtained in Experiment 4 (Mani & Johnson-Laird, 1982).

This experiment has replicated a number of findings obtained with the more elaborate or complex spatial descriptions used in Chapter 3 (Taylor & Tversky, 1992a; Tversky, 1991). One finding, however, has not been replicated. The spatial mental model account predicts that for these spatial descriptions the second item should be drawn first. While the second item was drawn first for most of the determinate descriptions and many of the indeterminate descriptions these differences were not significant. The overall probability of recall for the second item in a description was also higher than for any other item for both determinate and indeterminate descriptions. The evidence is thus consistent with the findings obtained in Chapter 4. One problem may be that exposure to reordered and inferable descriptions (in most cases with different first sentences) in the surprise recognition test (Experiment 4) may have introduced more error into the data. Another is that frequency may influence whether an item is drawn first



or second (even though there is no evidence that frequency influences other measures of recall for the determinate descriptions).

Overall the data support the hypothesis that people remember the structure of a spatial mental model and the process of its construction. The findings are complementary to the studies described in Chapter 3. In particular they support the conclusions of Experiment 4. However, they also confirm the suggestion that simple spatial descriptions of the type used here or by Johnson-Laird (1983) and Payne (1993) can vary considerably in memorability.



## 4.3 Experiment 6a

The purpose of this experiment was to investigate differences in difficulty in remembering different spatial descriptions of the kind used in this Chapter and in a number of other studies (Ehrlich & Johnson-Laird, 1982; Mani & Johnson-Laird, 1982; Oakhill & Johnson-Laird, 1984; Payne, 1993). A number of subjects mentioned that one description 'Landmarks' was particularly easy to visualize or picture and hence to remember subsequently. It is therefore reasonable to predict that measures of imagery values for items in the spatial descriptions may influence memorability.

This study was made up of two stages. In the first stage imagery, concreteness and meaningfulness values were obtained for the nouns contained in the eight spatial descriptions used earlier in this chapter. These values were then correlated with the number of correctly drawn descriptions in Experiment 5 and the verbatim and gist recognition scores obtained for each description in Experiment 4. In the second stage twenty-four new spatial descriptions (twelve determinate and twelve indeterminate) were generated. These spatial descriptions were then rated for difficulty by a sample of subjects from Experiment 4. The purpose of this was to select eight spatial descriptions which are neither particularly easy nor particularly difficult to remember for use in Experiment 6b.

### 4.3.1 Method

#### Subjects

Twenty-four rating questionnaires were sent out for stage two of the experiment. Eighteen people completed and returned the rating questionnaires. All the subjects had taken part in Experiment 4 or the pilot experiment for Experiment 4. Ten of the questionnaires were for indeterminate descriptions and eight for determinate descriptions.

#### Materials and procedure

The materials used in stage one of the experiment came from one of two sources. The eight spatial descriptions were taken from Experiment 4. The imagery norms were taken from those published by Paivio, Yuille and Madigan (1968). It should be noted that these norms were not produced by a British sample. However, all the nouns were common and were of everyday objects. In addition, there is evidence that the imagery norms do not differ greatly from those of British subjects (see Morris & Reid, 1972). Imagery norms were not available for all of the nouns in all of the descriptions (in fact for the 'Garden Tools' description no imagery norms were available). Where norms for several nouns were available the mean value was taken for the correlation.



<p><b>Determinate template 1</b></p> <p>The A is to the left of the B</p> <p>The C is to the right of the B</p> <p>The D is in front of the A</p> <p>The E is in front of the B</p>	<p><b>Determinate template 3</b></p> <p>The A is to the left of the B</p> <p>The B is behind the C</p> <p>The D is to the left of the C</p> <p>The E is in front of the C</p>
<p><b>Determinate template 2</b></p> <p>The A is behind the B</p> <p>The A is to the left of the C</p> <p>The A is to the right of the D</p> <p>The D is behind the E</p>	<p><b>Determinate template 4</b></p> <p>The A is in front of the B</p> <p>The A is behind the C</p> <p>The A is to the left of the D</p> <p>The E is behind the D</p>

**Figure 4.25: The four templates derived from the determinate descriptions in Experiment 4.**

The materials used in stage two were generated especially for this experiment. Twelve determinate descriptions were generated using the four determinate descriptions from Experiment 4 as templates (see Figure 4.25 for the four templates obtained). These were turned into descriptions by using twelve common categories which contained at least five relatively common objects. Objects were randomly assigned to each of the five serial positions in a description. Twelve indeterminate descriptions were generated by changing one spatial relation in each of the determinate descriptions. A rating questionnaire was constructed for the twelve determinate and the twelve indeterminate descriptions. People were asked to rate each of the twelve descriptions according to how difficult it was to understand the spatial relationships between objects in the description. They were provided with a rating scale from one ('very difficult') to seven ('very easy').

### 4.3.2 Results

**The relationship between imagery norms and memorability**

Values for imagery, concreteness and meaningfulness were correlated with the number of correctly drawn descriptions (Experiment 5) and the level of gist and verbatim recognition (Experiment 4) for seven of the eight spatial descriptions used earlier in this chapter. The subsequent correlations are depicted in Figure 4.26.



	Imagery	Concreteness	Meaningfulness
Correctly Recalled	-0.205	-0.667	0.727
Verbatim Recognition	-0.483	-0.107	-0.331
Gist Recognition	-0.128	-0.895	0.482

**Figure 4.26: Correlations between imagery values and memory scores.**

These figures were obtained using the Pearson product-moment correlation. The only positive correlations are between meaningfulness and gist recognition or drawing recall. The relationship between meaningfulness and the number of correctly drawn descriptions is significant ( $r = 0.727$ ,  $N = 7$ , one tailed  $p < 0.05$ ). There is also a significant negative relationship between concreteness and gist recognition ( $r = 0.895$ ,  $N = 7$ , one tailed  $p < 0.005$ ). However, it should be noted that small data sets ( $N = 7$ ) greatly reduces the power of the statistical test used.

### Ratings for the determinate descriptions

Ratings for the determinate descriptions were analyzed by ranking the ratings within subjects. Figure 4.27 shows the mean rank and standard deviation for all twelve of the determinate descriptions (in decreasing order of difficulty from top to bottom).

Determinate description	Mean Rank	Standard Deviation	Template
Transport	3.38	2.53	2
Stationery	4.56	1.86	2
Tools	4.81	3.08	3
Tableware	5.06	2.71	4
<i>Chess Pieces</i>	6.25	3.67	4
<i>Musical Inst.</i>	6.31	2.63	3
Creatures	6.31	1.80	4
<i>Clothing</i>	6.31	1.80	2
<i>Furniture</i>	6.50	3.46	3
Buildings	7.00	3.24	1
Fruit	8.44	2.20	1
Condiments	9.19	2.05	1

**Figure 4.27: Mean rank and standard deviation for the difficulty ratings of the determinate descriptions.**



Also indicated are the templates from which each description was generated. The determinate descriptions eventually used in Experiment 6b are in italics. The 'Creatures' description was not included in order to make it possible to select the indeterminate 'Creatures' description (see below). For the determinate descriptions there was clearly a marked difference in difficulty between items. There is some evidence that the way the description is phrased also influences perceived difficulty. Template 1 seems to be easiest, while template 2 seems to be hardest.

**Ratings for the indeterminate descriptions**

Rating for the indeterminate descriptions were analyzed in the same manner as the determinate descriptions. Figure 4.28 shows the mean rank and standard deviation for all twelve of the indeterminate descriptions (in decreasing order of difficulty from top to bottom). The range of perceived difficulty is smaller than that for the determinate descriptions. The variance within descriptions, though, appears to be greater. As with the determinate descriptions certain templates appear to produce more difficult descriptions than others. Template 2 appears to be relatively difficult and template 4 relatively easy (the templates for indeterminate descriptions differ from those of the determinate descriptions by a single altered spatial relation).

Indeterminate description	Mean Rank	Standard Deviation	Template
Transport	3.75	2.14	2
Condiments	5.06	1.79	1
Stationery	5.20	3.08	2
Clothing	5.81	2.97	2
Furniture	6.19	2.84	3
Musical Inst.	6.31	2.98	3
<i>Buildings</i>	6.38	2.93	1
<i>Fruit</i>	6.44	3.36	1
<i>Creatures</i>	6.88	2.93	4
Chess Pieces	6.93	4.46	4
<i>Tools</i>	7.00	2.83	3
Tableware	8.13	2.03	4

**Figure 4.28: Mean rank and standard deviation for the indeterminate descriptions.**



### 4.3.3 Discussion

When people are asked to rate simple determinate or indeterminate spatial descriptions the perceived difficulty of descriptions can vary considerably. For some descriptions these ratings are relatively consistent while others (especially some of the indeterminate descriptions) can produce very inconsistent ratings. The ratings obtained in this study were used to select ten spatial descriptions for Experiment 6b. The easiest determinate and the easiest indeterminate description were taken for use as practice items. Four determinate descriptions from the middle of the rankings were taken for the experiment proper. Four indeterminate descriptions were taken from the middle or the high end of the rankings to make certain that particularly difficult indeterminate descriptions were not used. Care was taken to avoid choosing indeterminate and determinate descriptions with the same category name to avoid confusion. One indeterminate description ('Chess Pieces') was excluded because of the very high variance among the ranks assigned to it.

Perhaps surprisingly, there is no evidence that imagery or concreteness (Paivio et al., 1968) were good predictors of recall or recognition for the descriptions used in Experiment 4. It is possible that the available norms are not sufficiently spread out to obtain reliable correlations because all the nouns used were concrete and high in imagery value. Also, the size of the sample may have contributed to the lack of a significant correlation (though it could not explain the apparent negative relationship between imagery and recall). There is some evidence to suggest that meaningfulness is related to recall for the spatial descriptions. Meaningfulness is measured by the number of associated words which can be written down for a given word within a certain time limit. It is possible that the advantage of well remembered descriptions like 'Landmarks' is that the different items in the description are higher in activation or association strength in long-term memory. Imagery may increase the association between elements in long-term memory when it relates two or more items together (Glenberg & Langston, 1992; Marschark, 1985; Marschark & Surian, 1989). The negative correlation between concreteness and gist recognition scores from Experiment 4 is also interesting. A similar negative correlation was obtained by Richardson (1987) between concrete versions of a three-term series task and reasoning performance. Richardson interpreted that finding as evidence that people utilized an amodal representation during reasoning. Presumably a very concrete representation interferes or competes with the amodal representations which people use to reason with (e.g. a concrete representation may require greater effort to construct, but provide no additional help in solving the three-term series problem). Such an interpretation would also be consistent with the abstract, structural, characteristics of a spatial mental model (see section 2.1.4), but not with the use of visual imagery. Of course other factors, such as the way a given description is structured, probably also influence the actual and perceived difficulty of a given description.



## 4.4 Experiment 6b

This experiment attempted to replicate and extend the findings obtained in Experiment 4. A number of important changes were made to the materials and procedures used in this study. The spatial descriptions used in the study were chosen from those rated in Experiment 6a in order to minimize differences in perceived difficulty between them. In Experiment 4 there were no differences in memory for descriptions that had been presented once or three times. In this study rather than manipulate the number of presentations it was decided to limit the amount of time a person could take with each description. To this end, mode of presentation was manipulated in order to investigate memory for a spatial mental model. The eight spatial descriptions used in the experiment were also divided into two sets. These two sets were counterbalanced for mode of presentation, order of presentation and for reordering in the phase 2 recognition task.

Half the spatial descriptions were presented, as before, in a written form. The remaining descriptions were read out loud to the subjects. It was predicted that people who heard the descriptions would have less time to elaborate their representation and hence tend to remember only the process of spatial mental model construction, for example in the form of an episodic construction trace (Payne, 1993). Descriptions which are read should provide more opportunity for elaboration and, as in Experiment 4, both the spatial mental model and the construction process should be remembered. Where both spatial mental model and episodic construction trace are remembered reordering should depress verbatim recognition more than gist. The inferable description has the same trace overlap as the reordered original and therefore the episodic construction trace should not support gist memory.

### 4.4.1 Method

#### Subjects

Twenty-seven members of staff or postgraduate students from the Open University volunteered to take part in the experiment. Each subject was offered two pounds in return for taking part. The age range of the subjects was from early twenties to mid forties.

#### Design

Two tasks were used in this experiment; a diagram matching task (phase 1) and a surprise recognition task (phase 2). There were three within subjects factors; mode of presentation (whether the spatial descriptions were *read* or *heard*), determinacy of the spatial descriptions (*determinate* or *indeterminate*) and the stability of the recognition items (whether the original description was *stable* or *reordered* in the recognition test). The two dependent variables were



*verbatim* recognition and *gist* recognition (see the results section for details of how these were scored). The eight spatial descriptions were randomly assigned to these eight conditions.

## Materials

Ten spatial descriptions were used in the experiment. They were constructed according to the procedures described by Payne (1993). Eight descriptions were chosen in order to minimize differences in memorability from a set of twenty-four items previously rated for difficulty (see Experiment 6a). The easiest determinate and easiest indeterminate description were chosen as the two practice items. The eight remaining spatial descriptions were divided into two sets (1 and 2) each containing two determinate and two indeterminate descriptions. Materials for the phase 1 diagram matching task were constructed slightly differently than in Experiment 4. One incorrect and one correct diagram was created for each description presented in the experiment. The correct diagrams depicted a configuration of five objects consistent with the four spatial relations of the original description. The incorrect diagram depicted the same array with one alteration. The alteration was consistent with three of the four spatial relations of the original spatial descriptions but violated the remaining spatial relation.

For phase 2 two sets of recognition materials were prepared. For set A half the determinate and half the original indeterminate spatial descriptions were randomly selected and reordered. In set B the reordered items from set A were replaced with the stable original descriptions. The remaining original descriptions were then reordered. As in Experiment 4 the recognition materials consisted of an original description (stable or reordered), an inferable description and two foils. The reordered original, inferable and two foils were constructed in the same way as those in Experiment 4 (Mani & Johnson-Laird, 1982; Payne, 1993). Each recognition item (apart from the stable or reordered originals) contained one sentence in common with the original description. A trace overlap of one was preserved between the original description and the reordered original, inferable and the two foils.

## Procedure

The same general procedure as that in Experiment 4 was adopted (Mani & Johnson-Laird, 1982; Payne, 1993). Slight changes were made to the instructions for phase 1. Before carrying out the two practice trials subjects were given a sheet of paper explaining how the diagrams for the phase 1 matching task were depicted. In particular it was indicated that spatial relations such as 'in front of' did not include 'diagonally in front of'. It was also indicated that where a given spatial relation between two objects was used it was possible for there to be an intervening object. For the practice items and for the initial set of descriptions in the diagram matching task (set 1 or set 2) the mode of presentation was kept constant (the descriptions were either read or heard). For the second set of descriptions (either set 1 or set 2) a different mode of presentation (read or



heard) was used. Where descriptions were read the subject was presented with a card containing the description and asked to indicate when he or she was ready to proceed with the diagram matching task. Where descriptions were heard the experimenter read each description twice, in order. The experimenter paused at the end of each sentence and when the subject indicated they were ready went on and read the next sentence. In the practice trials and in the main experiment each subject saw two diagrams for each spatial description. One diagram was correct and one incorrect. The two diagrams were presented in a random order immediately after the appropriate original spatial description had been read or heard. Subjects were asked to respond 'good match' or 'bad match' for each of the ten spatial descriptions presented in the practice task and the rest of phase 1.

The procedure for the phase 2 recognition test also differed only slightly from that in Experiment 4. Each subject was presented with one of two sets of recognition materials (set A or set B) to ensure that each original description was presented in both stable and reordered forms during the recognition task. Before phase 2 the recognition materials were shuffled to ensure that the items in the recognition task was completed in a random order. Mode of presentation, the order of descriptions sets (1 or 2) in phase 1 and the recognition materials in phase 2 (set A or B) were counterbalanced between subjects. This and the random ordering of descriptions within phase 1 and 2 controlled for consistent time differences between learning and test conditions.

#### **4.4.2 Results**

##### **Performance on the diagram matching task**

Overall performance (79%) on the diagram matching task was lower than that of subjects in Experiment 4. Out of twenty-seven subjects four people scored lower than fifty percent when matching either the determinate descriptions, the indeterminate descriptions or both. These four subjects were excluded from the remaining analyses.

##### **Overall recognition performance**

Mean recognition scores are summarized by description type in Figure 4.29 below. As before chance performance is 25% for verbatim recognition and 16.7% for gist recognition. The pattern of results is relatively complex. For this reason, levels of gist and verbatim recognition performance are analyzed in more detail below.



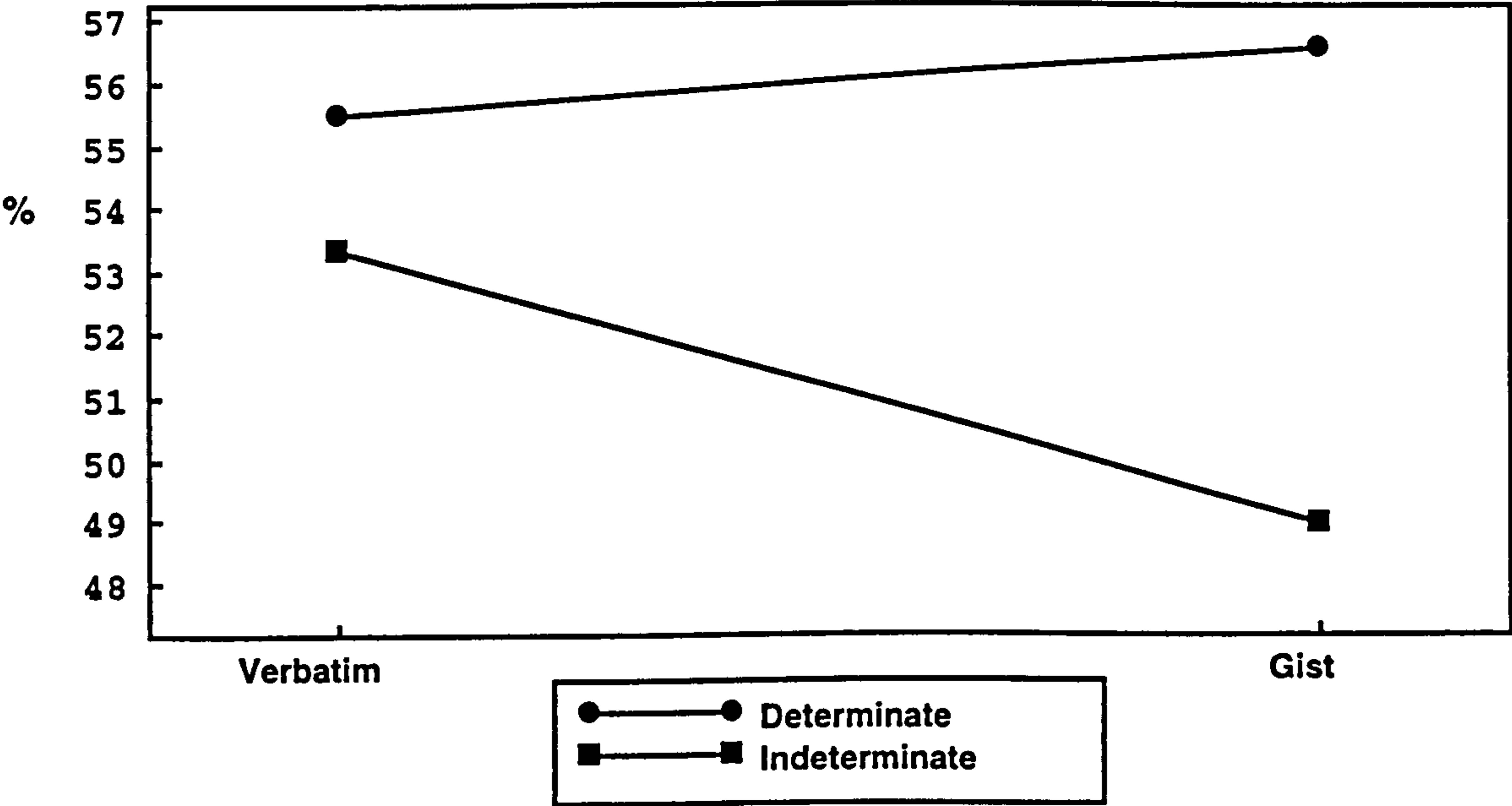
<u>Description type</u>		<u>Verbatim</u>	<u>Gist</u>
Determinate	Stable	58.7 %	73.9 %
	Reordered	58.7 %	60.9 %
Indeterminate	Stable	73.9 %	50.0 %
	Reordered	30.4 %	34.8 %

**Figure 4.29: Mean percentage verbatim and gist recognition by description type**

*Is there an interaction between determinacy and type of recognition memory?*

Overall recognition performance showed a pattern consistent with memory for a spatial mental model; gist memory was higher for determinate descriptions and verbatim memory was higher for indeterminate descriptions. Figure 4.30 shows the interaction between determinacy and gist or verbatim recognition.

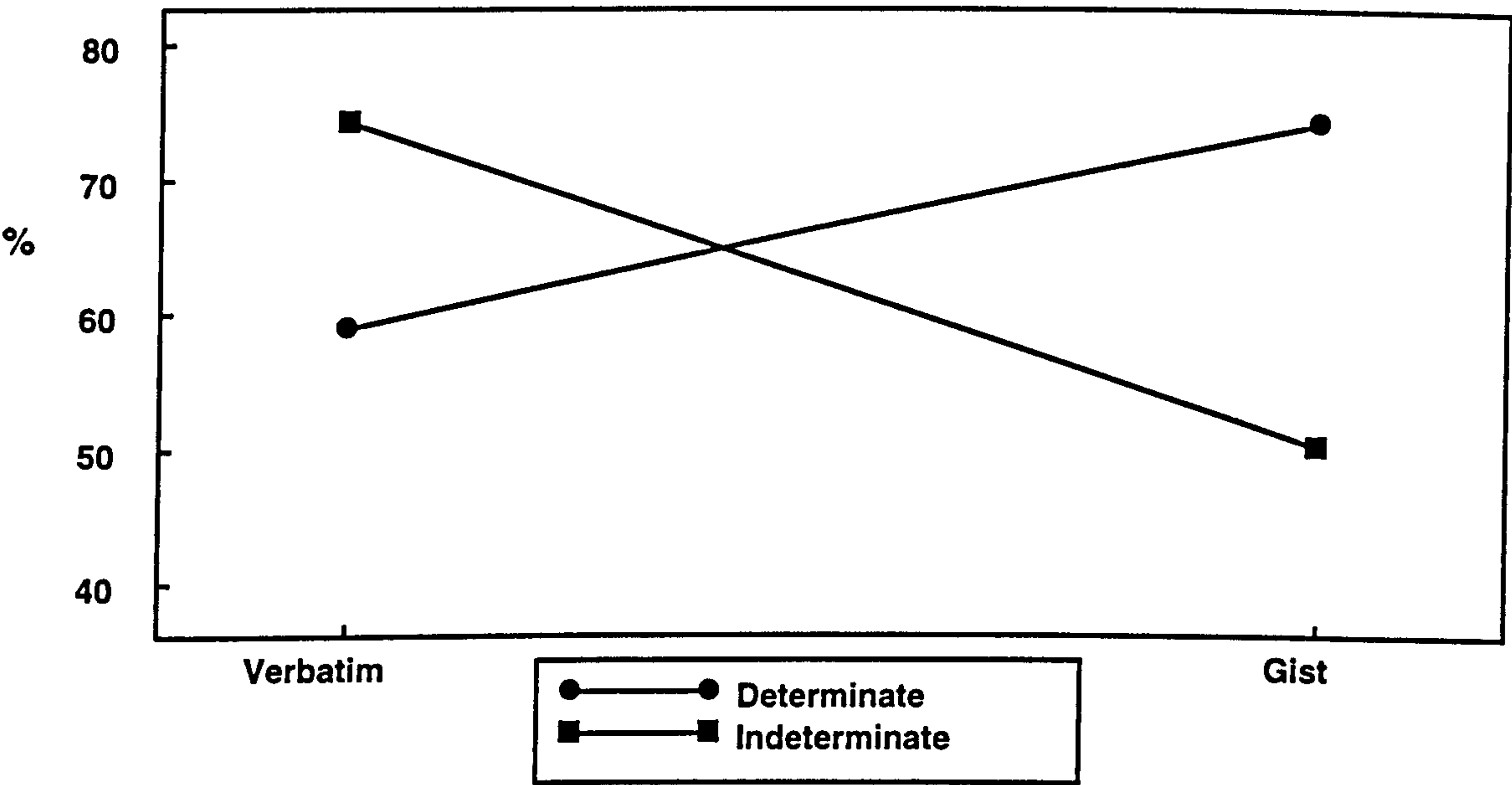
However, this analysis gives a misleading picture of recognition performance for determinate and indeterminate descriptions. Close inspection of Figure 4.29 above will reveal that memory for the reordered indeterminate descriptions is barely above chance. It is possible that this floor effect is partially concealing an interaction between determinacy and type of recognition memory. In addition to the overall analysis reported above recognition performance was analyzed separately for stable and reordered descriptions.



**Figure 4.30: Mean percentage gist and verbatim recognition by type of description.**



Figure 4.31 below depicts the crossover interaction for stable descriptions only. The significance of this interaction was tested using a chi square design (refer to Experiment 4 for a detailed discussion of the scoring and analysis of the interaction predicted by Mani and Johnson-Laird).



**Figure 4.31: Mean percentage gist and verbatim recognition by type of description (stable descriptions only).**

The analysis compares the frequency with which determinate and indeterminate descriptions supported gist *or* verbatim recognition (as before, data points were pooled across subjects as well as across descriptions). The contingency table for the chi square test is shown in Figure 4.32 below:

		Type of recognition:	
		Verbatim	Gist
Type of description:	Determinate	4	11
	Indeterminate	13	3

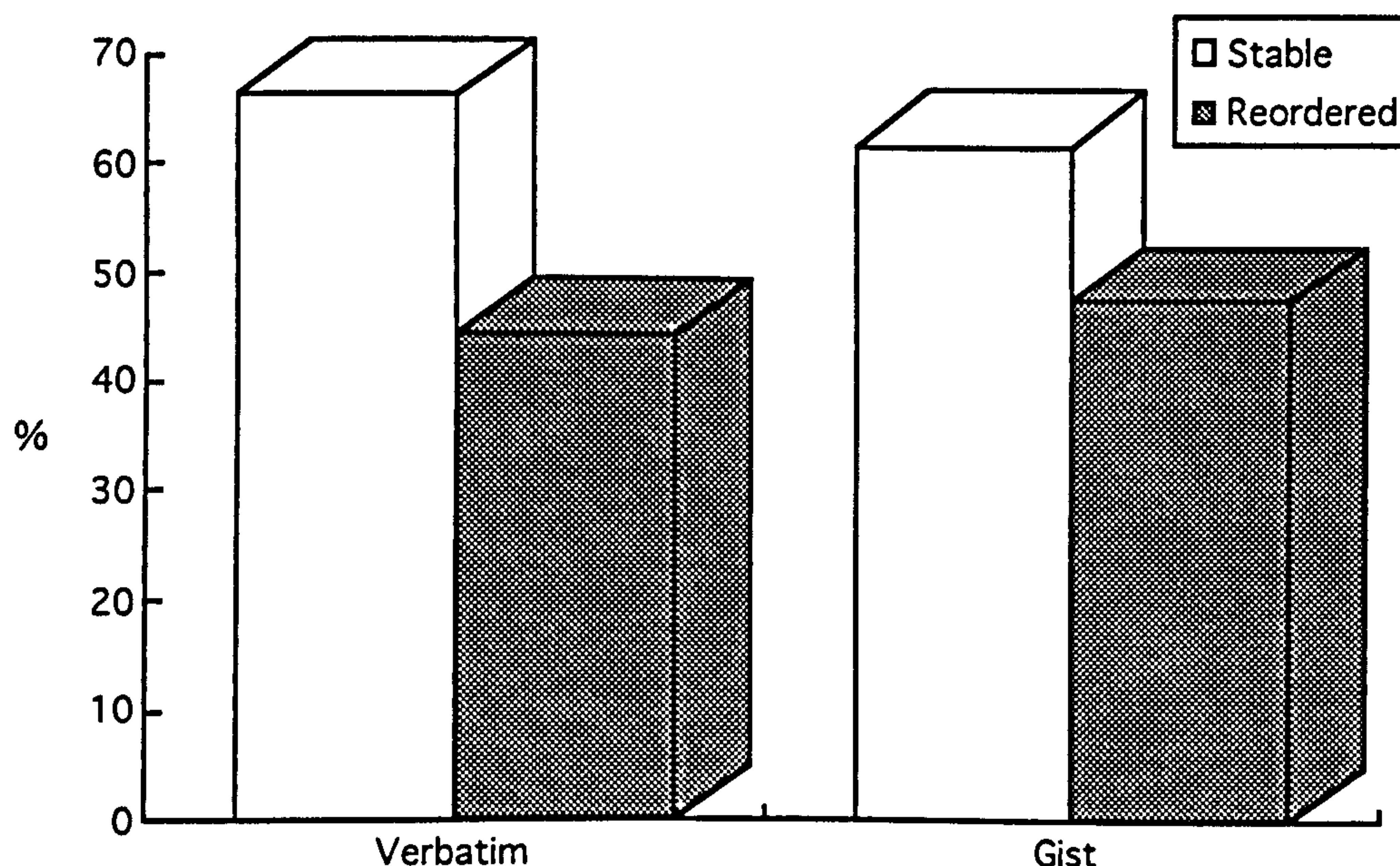
**Figure 4.32: The frequency of verbatim *or* gist recognition for determinate and indeterminate descriptions (stable descriptions only).**

Because of the low frequencies in two of the cells these data were analyzed with the Fisher Exact test. This confirmed that this pattern of cell frequencies was significant (N = 31, one tailed p = 0.0031). This replicates the finding from Experiment 4 that gist recognition scores are higher for determinate descriptions and verbatim recognition scores are higher for indeterminate descriptions.



### *Does reordering depress recognition memory?*

The overall effect of reordering on memory was analyzed separately for gist and verbatim recognition. Figure 4.33 shows mean percentage gist and verbatim recognition for stable and reordered descriptions. Differences were tested using the Wilcoxon signed-ranks test.



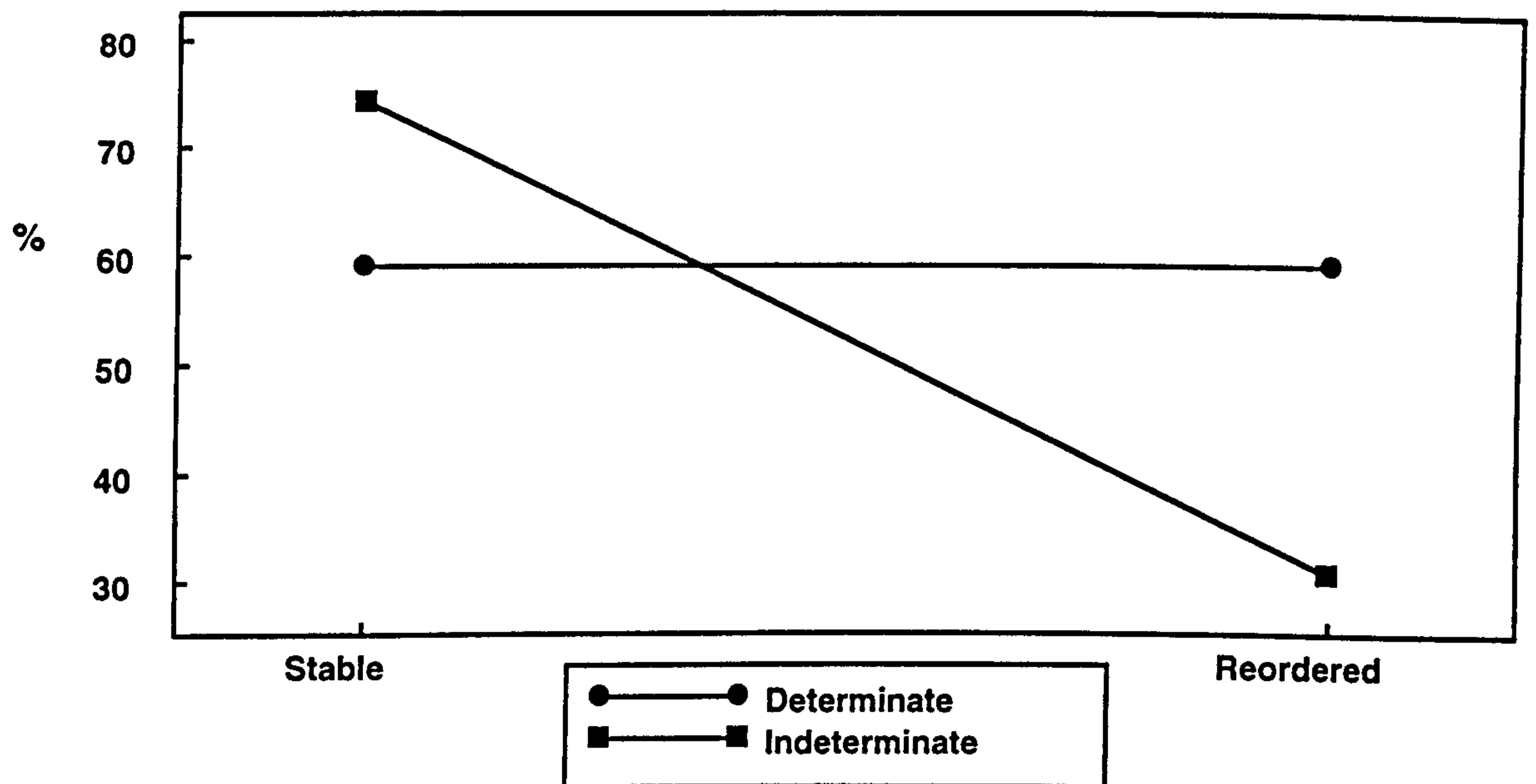
**Figure 4.33: Mean percentage gist and verbatim recognition for stable and reordered descriptions.**

This was significant for both verbatim ( $N = 23$ ,  $Z = 2.55$ , two tailed  $p = 0.011$ ) and for gist recognition ( $N = 23$ ,  $Z = 2.12$ , two-tailed  $p = 0.034$ ). There is also evidence of an interaction between the stability of the recognition test item and determinacy. This interaction is investigated further in the next section.

### *The effect of reordering on determinate and indeterminate descriptions*

Examination of Figure 4.29 above shows that reordering appears to depress verbatim recognition virtually to chance levels for the indeterminate descriptions and to have little or no effect on determinate verbatim descriptions. This pattern, however, does not occur for gist recognition suggesting that there is a three-way interaction between determinacy, reordering and type of recognition. Because of the difficulties of analyzing three-way interactions using non-parametric methods analyses of two-way interactions were carried out separately for gist and verbatim recognition scores. This was done using post-hoc non-parametric factorial analysis of variance (Meddis, 1984). Figure 4.34 shows the interaction between determinacy and reordering on verbatim recognition performance.

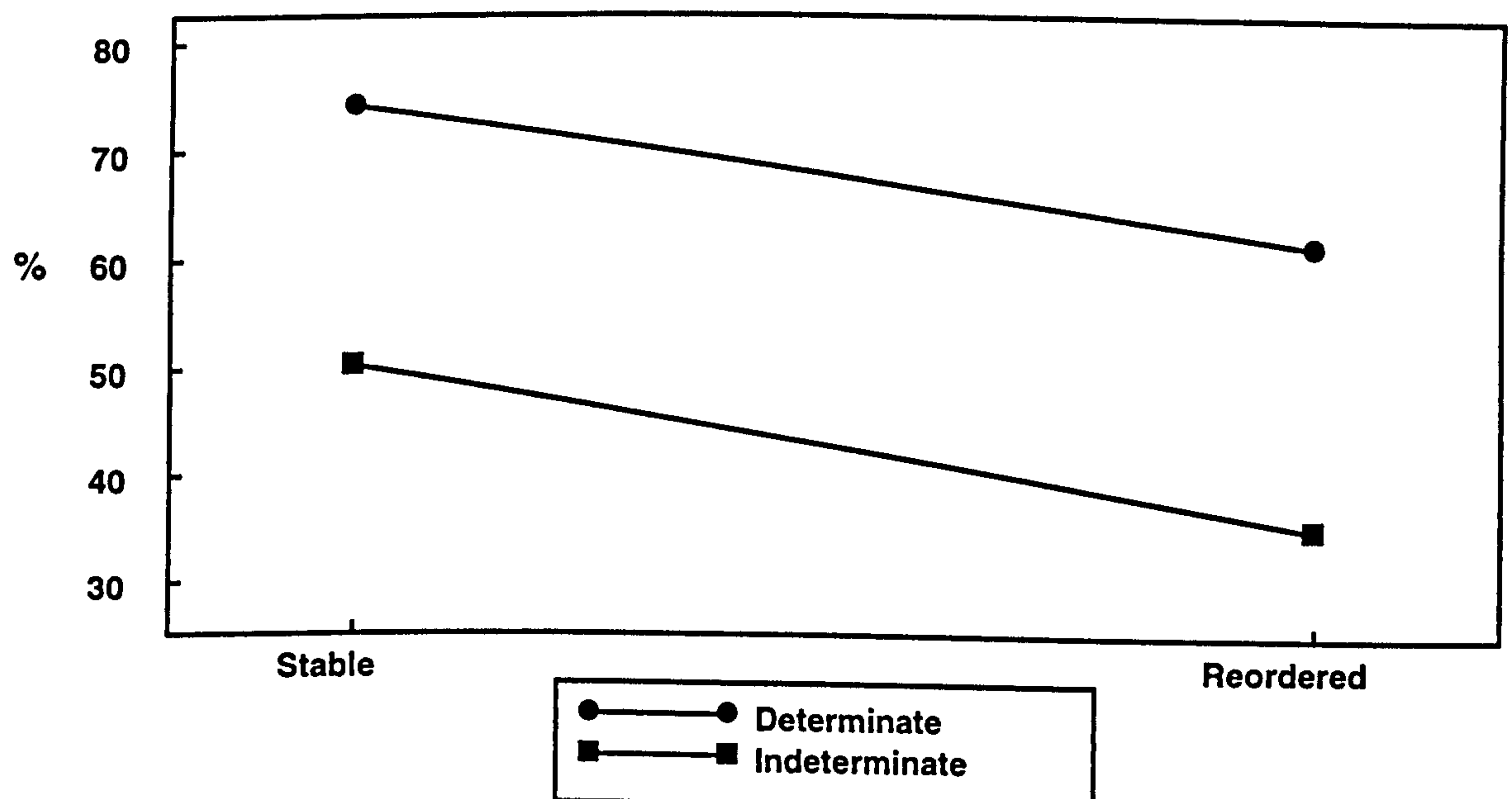




**Figure 4.34: The effects of reordering and determinacy on verbatim recognition.**

This interaction on verbatim recognition performance was significant on a post-hoc test for an interaction ( $H = 7.79$ ,  $N = 23$ ,  $H_{crit} = 6.25$ , one tailed  $p < 0.05$ ).

Figure 4.35 depicts the level of determinate and indeterminate gist recognition scores for reordered and stable descriptions. There is a significant main effect of determinacy ( $H = 12.54$ ,  $N = 23$ ,  $H_{crit} = 7.79$ ,  $p < 0.01$ ) and of reordering on gist recognition (see previous section). However, there is no evidence of a significant interaction between the two ( $H = 0.096$ ,  $N = 23$ ,  $H_{crit} = 7.79$ ,  $p = n.s.$ ).



**Figure 4.35: The effects of reordering and determinacy on gist recognition.**



### *Recognition memory for heard descriptions*

Figure 4.36 shows the pattern of verbatim and gist recognition for the different types of spatial descriptions that were heard by subjects. The recognition scores are similar to those for overall recognition (see Figure 4.29 above). There is one important difference. Unlike the overall recognition scores reordering depresses verbatim and gist memory for both determinate and indeterminate descriptions.

<u>Description type</u>		<u>Verbatim</u>	<u>Gist</u>
Determinate	Stable	69.6 %	82.6 %
	Reordered	60.9 %	60.7 %
Indeterminate	Stable	78.3 %	52.2 %
	Reordered	34.8 %	30.4 %

**Figure 4.36: Mean percentage verbatim and gist recognition by description type (heard descriptions only).**

Recognition for the reordered indeterminate descriptions is barely above chance. For this reason recognition memory for the reordered descriptions was not included in the analysis of the interaction between determinacy and type of recall. The crossover interaction predicted by the spatial mental model account (high gist memory for determinate descriptions and high verbatim memory for the indeterminate descriptions) is present for the stable descriptions. This was tested using a chi square design (see above). This interaction was confirmed as significant for the stable descriptions using the Fisher Exact test ( $N = 17$ , one-tailed  $p = 0.036$ ).

However, unlike the overall scores verbatim memory does not show a significant interaction between determinacy and reordering ( $L = 16$ ,  $N = 23$ ,  $Z = 1.63$ , two-tailed  $p = 0.10$ ). The effect of reordering to reduce verbatim recognition ( $N = 23$ ,  $Z = 2.72$ , one-tailed  $p = 0.033$ ) and gist recognition ( $N = 23$ ,  $Z = 2.27$ , one-tailed  $p = 0.012$ ) is significant with the Wilcoxon signed-ranks test. The three-way interaction present in the overall analysis is not indicated in recognition performance when the descriptions are heard.

### *Recognition memory for read descriptions*

Recognition memory for the descriptions read by subjects also shows similarities with overall recognition scores (see Figure 4.37 below and Figure 4.36 above). Again, however, it differs in one major respect. Verbatim recognition for reordered determinate descriptions is actually superior to that for stable determinate descriptions. This interaction between determinacy and reordering is significant for verbatim recognition ( $L = 24$ ,  $N = 23$ ,  $Z = 2.45$ , two-tailed  $p = 0.014$ ) but not for gist



recognition. Therefore it appears that the three-way interaction between determinacy, type of recall and reordering found in the overall analysis is present in performance for read descriptions but not for heard descriptions (refer to Figures 4.29, 4.36 and 4.37).

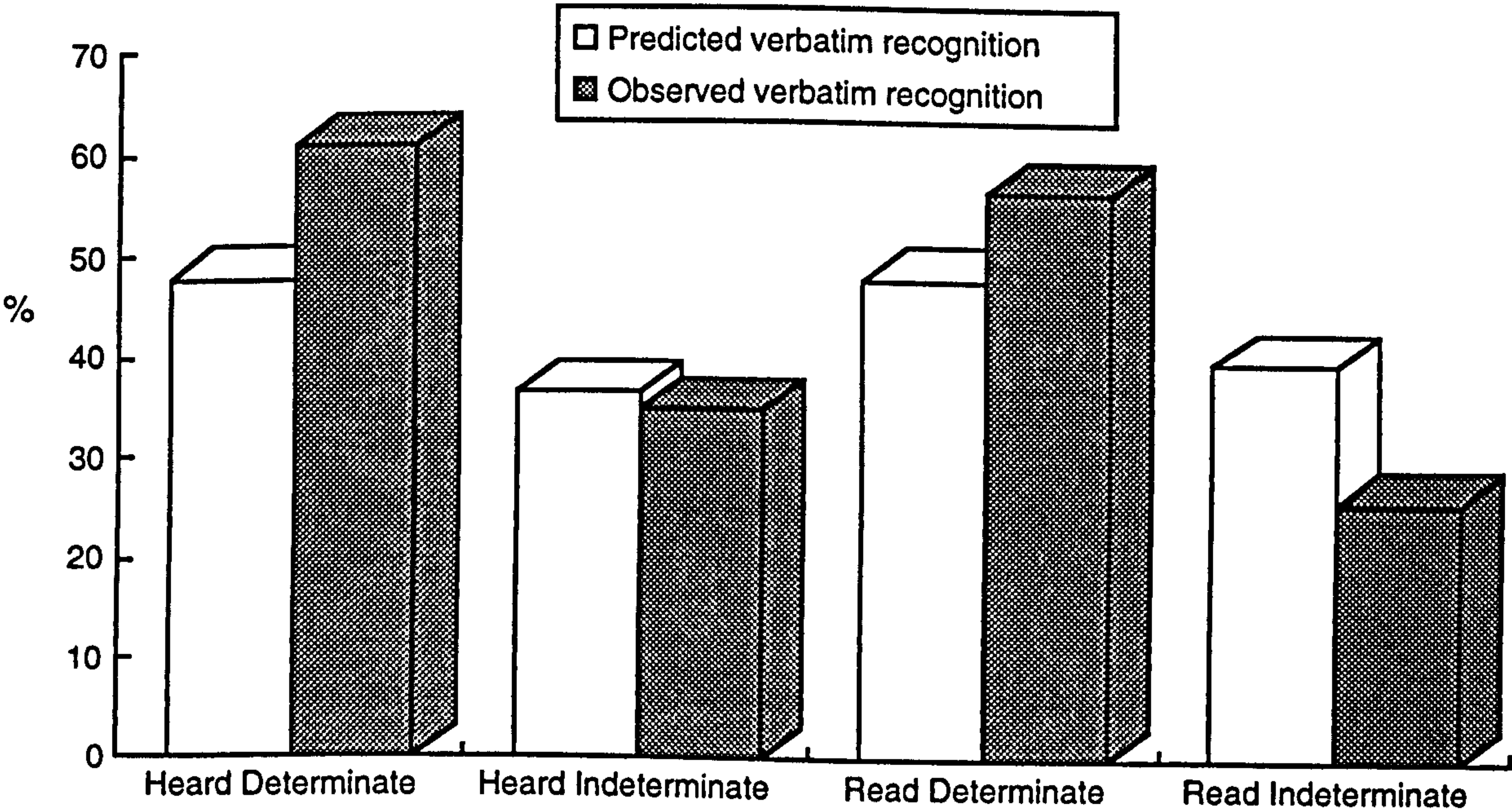
<u>Description type</u>		<u>Verbatim</u>	<u>Gist</u>
<b>Determinate</b>	<b>Stable</b>	47.8 %	65.2 %
	<b>Reordered</b>	56.5 %	60.9 %
<b>Indeterminate</b>	<b>Stable</b>	69.6 %	47.8 %
	<b>Reordered</b>	26.1 %	39.1 %

**Figure 4.37: Mean percentage verbatim and gist recognition by description type (read descriptions only).**

The effect of reordering on recognition was analyzed using the Wilcoxon signed-ranks test. For read descriptions reordering does not depress gist recognition (N = 23, Z = 0.85, two-tailed p = 0.40) or verbatim recognition significantly (N = 23, Z = 1.56, two-tailed p = 0.12). Only for read indeterminate descriptions does reordering significantly depress verbatim recognition (N = 23, Z = 2.53, two-tailed p = 0.023).

*Memory for the language of the description*

When levels of verbatim recognition for reordered descriptions are above chance this does not necessarily imply that the language of the description has been remembered. However it is possible to estimate by how much verbatim memory exceeds chance under certain circumstances.



**Figure 4.38: Observed and predicted verbatim recognition for reordered descriptions.**



This is because where gist recognition for reordered descriptions exceeds verbatim recognition (as in this experiment) it indicates memory for a spatial mental model. Where gist recognition is present verbatim memory should be scored 50% of the time by chance. Where gist recognition is not present verbatim memory arises 25% of the time by chance. Figure 4.38 shows the levels of verbatim recognition predicted by chance and those actually observed for reordered descriptions only. Only for determinate descriptions does verbatim recognition for reordered descriptions exceed that which can be accounted for by the observed level of gist recognition.

### 4.4.3 Discussion

Constructing spatial mental models of spatial descriptions in working memory can result in the retention of three different kinds of memory trace; memory for the language of the description, memory for the construction of a spatial mental model or memory for the structure of a spatial mental model. In this experiment people who are presented with spatial descriptions are able to remember the structure of a spatial mental model. For determinate descriptions recognizing the structure of a spatial mental model is easier than recognizing the language of the original description. For indeterminate descriptions (where constructing a spatial mental model is difficult if not impossible) recognizing the language of the original description is easier than recognizing the spatial situation it describes. This crossover interaction is difficult to explain without proposing at least two different kinds of representation (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982; Payne, 1993). One representation preserves the structure of the described situation, the other representation preserves information about the spatial description. The interaction is obtained even when descriptions are controlled for perceived difficulty, items on the recognition test are controlled for trace overlap or when the descriptions are heard rather than read.

The effects of reordering on memory for the spatial descriptions in this experiment differ according to the determinacy of the description, whether gist or verbatim memory is analyzed and whether the descriptions are read or heard. For indeterminate descriptions reordering disrupts verbatim recognition almost to chance. In fact if gist recognition levels are taken into account verbatim recognition for reordered indeterminate descriptions is below chance in every case.<sup>3</sup> When descriptions are heard, reordering significantly depresses both gist and verbatim recognition memory for all descriptions (though only indeterminate verbatim recognition is

---

<sup>3</sup> On average gist recognition should produce verbatim recognition fifty percent of the time. Chance verbatim recognition is 25%. The overall reordered indeterminate gist recognition is 34.8%. This predicts that verbatim recognition should be at  $(34.8 * 0.5) + ((100-34.8) * 0.25) = 33.7\%$ . The actual level is 30.4%.



reduced to chance levels). When descriptions are read, reordering significantly depresses recognition memory for indeterminate descriptions. However, determinate descriptions that were read by subjects show no significant reductions in either verbatim or gist recall.

These results suggest that memory for indeterminate descriptions (whether they are read or heard) is almost entirely based on a memory trace that is disrupted by reordering. This cannot be memory for the structure of a spatial mental model because a reordered description still describes the same situation. Nor can it be memory for the language of a description because reordering leaves the propositions of a description intact. However, a representation of the process and hence order of spatial mental model construction such as the episodic construction trace is disrupted by reordering the original description (Payne, 1993). The interaction between determinacy and type of recall indicates that memory for determinate descriptions appears to be based primarily on memory for the structure of the situation referred to by the description (Mani & Johnson-Laird, 1982). However, when a description is heard reordering also depresses gist and verbatim recognition for determinate descriptions. When a description is read reordering has little or no effect on recognition for determinate descriptions.

These findings fit well with the spatial mental model account outlined in Chapter 2. When people read or hear descriptions they attempt to construct a spatial mental model. For determinate descriptions the spatial mental model they construct is coherent and complete. For indeterminate descriptions the spatial mental model they attempt to construct is often inconsistent or incomplete. At a later stage they are given a recognition test. For the determinate descriptions they are able to identify those descriptions which are consistent with their memory of the spatial mental model they constructed. For the indeterminate descriptions few people will have been able to cope with the indeterminacy of the description. As they read the alternatives in the recognition test they attempt to construct spatial mental models of them. Where an original description has not been reordered the attempt to construct a spatial mental model will remind them of their previous attempt to construct a spatial mental model. Where the original description has been reordered the attempt to construct a spatial mental model during recognition will not remind them of their previous attempt to construct a spatial mental model.

The different effect of reordering on determinate recognition depending on whether a description is read or heard is harder to explain. It cannot be the case that reading determinate descriptions provides more opportunity for an elaborate spatial mental model to be constructed; recognition memory for read descriptions is slightly worse than that when descriptions are heard. Also there is no obvious reason why good memory for a spatial mental model should prevent people remembering the process of its construction. Perhaps the opposite, that hearing a description results in a more elaborate spatial mental model is the case. Unfortunately this



explanation is problematic because it suggests that the construction of a spatial mental model is less likely to be remembered than the mental model itself. This is contradicted by recognition performance for indeterminate descriptions which appears to consist almost entirely of memory for an episodic construction trace. An alternative explanation rests on the possibility that there is something about hearing a description that promotes retention of an episodic construction trace or something about reading a description that interferes with it.

Hearing a description ensures that each sentence is presented in order, while when a description is read some sentences may be read out of sequence. When hearing a description a person may concentrate harder on their attempt to construct a spatial mental model; they cannot deviate from the original order in any way. On rereading a description people may skip sentences that describe spatial relations they are sure about, they may even reread the whole description out of order e.g. in reverse. It is not suggested that people set out to read the description out of sequence but that once they have established the spatial relations between objects in a description they may rehearse, refresh or check their spatial mental model by reading sentences out of sequence. This may result in a less coherent episodic construction trace. The process of rereading may also improve memory for the language of the text, though there is no direct evidence for this. These differences may not be marked for indeterminate descriptions because of the difficulty in constructing a spatial mental model. When reading an indeterminate spatial description people may prefer not to deviate from the original order because the indeterminacy only arises later in the description (either in the second, third or fourth sentence). Thus hearing a spatial description may promote the retention of a single coherent episodic construction trace. Reading a spatial description may enable some readers to deviate from the spatial mental model construction order as they refresh or check the spatial mental model they have constructed. An alternative explanation is that reading, but not hearing, a description interferes with the spatial mental model construction process. However, this explanation does not explain why reading an indeterminate description would not interfere with spatial mental model construction.

In this experiment people who read or heard spatial descriptions comprehend them by attempting to construct spatial mental models consistent with the information contained in the descriptions. If the spatial description was determinate they appear to remember the structure of the spatial mental model they constructed. They also show some evidence of memory for the language of the description. If the description was indeterminate they remember their attempt to construct a spatial mental model. They also show some evidence of memory for a spatial mental model of the indeterminate description, but no evidence of memory for the language of the description. For determinate descriptions there is no direct evidence that retention of an episodic construction trace precedes retention of the structure of a spatial mental model. The level of recognition memory for indeterminate descriptions suggests that retention of an episodic construction is possible even when little or no memory for a spatial mental model is observed.



## 4.5 General Discussion

The three experiments reported in this chapter investigate the recognition and recall of short descriptions containing the spatial relations between five objects. It has been argued that, in each case, the findings support the account of spatial mental models outlined in Chapter 2. In that account three ways of remembering a spatial description were proposed. The following three sections briefly discuss the evidence in favour of each of the accounts. The predictions made by the three accounts are summarized in Figure 4.4 (repeated here for convenience).

<u>What is remembered:</u>	<u>Prediction:</u>
The language of the description (comprehension processes)	i) No effect of reordering. ii) Verbatim recognition higher than gist.
An episodic construction trace (construction processes)	i) Reordering reduces recognition. ii) Verbatim recognition higher than gist.
The structure of a spatial mental model (consultation processes)	i) High determinate gist recognition. ii) High indeterminate verbatim recognition.

**Figure 4.4:** The predictions from three different accounts of memory for spatial descriptions.

### 4.5.1 Memory for the language of a description

In both Experiment 4 and 6b gist recognition scores are significantly higher than verbatim recognition scores for the determinate descriptions. This finding is inconsistent with an account which proposes that people remember *only* the sentences of the description (or a propositional representation based on those sentences). Further evidence against the view that people remember the language of the description is given by the positive correlation between backward digit span and gist recognition scores. If people learn and remember the sentences of the description measures of verbal short-term memory such as backward digit span should be correlated with verbatim recognition scores. In addition very few subjects reported remembering or trying to remember the language of the descriptions. The evidence suggests that remembering the sentences of a description is not a deliberate strategy adopted by many subjects.



While there is little evidence that remembering determinate descriptions involves memory for the language of the descriptions, this may not be the case for indeterminate descriptions. Memory for the language of the descriptions presented in Experiment 4 and 6b is indicated by the level of verbatim recognition for the reordered descriptions. The reordered descriptions exclude the possibility that verbatim recognition may be based on memory for an episodic construction trace (this point is addressed more fully in section 4.5.2 below). In Experiment 4 verbatim recognition scores slightly exceed gist recognition scores for the reordered indeterminate descriptions, though this difference is not significant (see Figure 4.6). In Experiment 5 the probability of recalling an object from an indeterminate description is predicted by the number of times that an object was mentioned in the original description. This is consistent with a propositional representation where objects mentioned in more than proposition are more likely to be recalled. Unfortunately, such a finding is consistent with memory for the sentences of the description or with an episodic construction trace.

A more serious problem for any account which proposes that spatial descriptions are retained in a form close to that of the original language is that the sentences of indeterminate descriptions are not always remembered. In Experiment 6b verbatim recognition scores for the reordered indeterminate descriptions are barely above chance. This finding is a serious blow even to an account of memory for spatial descriptions which argues that the language of the description is retained only if a spatial mental model can not be constructed. It also rules out the Mani and Johnson-Laird (1982) explanation of why verbatim recognition scores exceed gist recognition scores for indeterminate descriptions. It is not the case that people *always* switch to remembering the sentences of a description when spatial mental model construction is difficult or impossible. A similar finding has been observed by Gernsbacher (1990; 1991) who has shown that scrambling a story (which would make constructing a mental model more difficult) does not necessarily improve memory for the language of the story. A more plausible explanation is that people continue to attempt construction of a spatial mental model. Such an explanation could explain high verbatim recognition in the stable indeterminate descriptions in terms of people remembering their (probably unsuccessful) attempt to construct a spatial mental model. In conclusion, there is little evidence that people remember the language of the descriptions they have read.

#### 4.5.2 Memory for an episodic construction trace

Memory for an episodic construction trace can be demonstrated if reordering a description impairs memory for that description in a recognition test. Reordering impairs recognition memory because it reduces the overlap between the operations used to construct the original, stable description and those used to construct the reordered description. Evidence for the effect of reordering is particularly strong when the original description was heard by subjects (as for half the



descriptions in Experiment 6b). This is consistent with the fact that when a description is heard it is virtually impossible for a listener to deviate from the order the description is presented in. When the description is written down it is possible for subjects to read sentences out of the original sequence (as they may have done for some descriptions in Experiment 6b). For the descriptions which were read by subjects the effect of reordering is significant for only some of the descriptions. In Experiment 4 only the matching descriptions showed an effect of reordering (this may have been due to differences in difficulty for the non-matching descriptions). In Experiment 6b only the indeterminate description read by the subjects showed a significant effect of reordering. Taken together these results support the conclusion that people do remember the process of spatial mental model construction, and that this memory may take the form of an episodic construction trace (Payne, 1993).

However, the results obtained in Experiment 5 are problematic for the episodic construction trace hypothesis (though not necessarily for the view that people remember the process of spatial mental model construction). As predicted, objects and spatial relations from determinate descriptions are more likely to be drawn (and to be drawn earlier) if they are integrated into a spatial mental model early in its construction. The problem for the episodic construction trace is that this does not always appear to be the case for the indeterminate descriptions. Indeterminate descriptions show only a weak relationship between the probability of recalling an object and the order of spatial mental model construction. By contrast number of times an object was mentioned in the original description is a significant predictor of recall from an indeterminate description. This result suggests that recall for indeterminate, but not determinate descriptions, may be based on a propositional representation such as an episodic construction trace or memory for the sentences of the description (refer to section 4.5.1 above).

The differences between memory for the indeterminate and determinate descriptions can not be accounted for by the current version of the episodic construction trace hypothesis. This is because the current version of the episodic construction trace hypothesis records operations in much the same way for both determinate and indeterminate descriptions. If the episodic construction trace is the source of all order effects in recall obtained in this thesis (like those discussed in Experiment 5 and those in Chapter 3) then those order effects should be obtained for both determinate and indeterminate descriptions. The picture is complicated slightly by the fact that drawing orders for the indeterminate descriptions do appear to show an above chance relationship with the predicted order of spatial mental model construction. Indeterminate descriptions also show a significant relationship between memory for spatial relations and predicted spatial mental model construction order. One possible interpretation is that determinate descriptions result in a stronger episodic construction trace, but this explanation is inconsistent with the evidence that indeterminate descriptions are influenced more strongly than determinate descriptions by reordering. The conclusion favoured here is that memory for spatial mental model construction order is a result of a combination of both *construction* and *consultation*



processes (see 2.5.1). The basis of this interpretation is that determinate descriptions differ from indeterminate descriptions only in the extent to which it is possible to complete spatial mental model construction for them. A more detailed version of this explanation is considered in Chapter 6 (see sections 6.1.2 and 6.1.5).

### 4.5.3 Memory for the structure of a spatial mental model

Memory for the structure of a spatial mental model is indicated by gist recognition where subjects 'recognize' an inferable description in a recognition test. The inferable description describes the same arrangement of objects as the original description subjects were presented with. In Experiments 4 and 6b the inferable description shares only one sentence with the original description, and only one of the operations used to construct a spatial mental model from the inferable description overlaps with the operations used to construct a spatial mental model from the original description. In both experiments gist recognition scores are higher than verbatim recognition scores where the description supports the construction of a single, coherent spatial mental model (i.e. where the description is determinate). This pattern is reversed for indeterminate description where more than one spatial mental model can be constructed.

This interaction is very difficult to explain except by appealing to two kinds of representation (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). The first kind of representation, a spatial mental model, preserves the structure of the situation conveyed by a description. The second kind of representation must support verbatim recognition memory for the description, though it need not preserve the language of the description (see 4.5.1 and 4.5.2 above). Further support for this distinction was obtained in Experiment 5. The order and probability of recall in a drawing task was significantly related to the order of spatial mental model construction, but only for the determinate descriptions. For the indeterminate descriptions the number of times an object was mentioned in a description was a better predictor of recall than spatial mental model construction order. Again, these findings are consistent with the view that people construct and remember spatial mental models when they are presented with determinate descriptions.

## 4.6 Conclusions

In the introduction to this chapter a hypothesis about the time course of memory for spatial descriptions was proposed. No direct evidence was obtained to suggest that the structure of a spatial mental model is only retained after an episodic construction trace has been formed. However, poor recognition for reordered indeterminate descriptions in the final experiment suggests that it is possible to retain an episodic construction trace without remembering the



structure of a spatial mental model, provided spatial mental model construction is difficult or incomplete. Two experiments have shown that people are able to remember both an episodic construction trace (Payne, 1993) and the structure of a spatial mental model (Mani & Johnson-Laird, 1982). A third experiment demonstrated that drawing order and the probability of recall from determinate spatial descriptions can be predicted by the order of spatial mental model construction. This supports the conclusions drawn in Chapter 3 that drawing a remembered scene involves reconstructing a spatial mental model (Taylor & Tversky, 1992a; Tversky, 1991) and that items integrated into a spatial mental model early in its construction are remembered better (Denhière & Denis, 1988; Gernsbacher, 1991).



## Chapter 5 - The long-term retention of spatial information

This chapter investigates memory for spatial information in the form of maps, routes and verbal descriptions at retention intervals of months rather than minutes or hours. Many studies of very long-term memory have focused on knowledge gained during high school or university education (Bahrick, 1983; 1984; Cohen, 1989; Cohen, Conway, & Stanhope, 1992). A number of studies have investigated the retention of spatial information over intervals of months or years but these have focused almost exclusively on knowledge gained by navigation through the environment, typically a university campus (Bahrick, 1983; Herman, Cachuela, & Heins, 1987; Herman, Kail, & Siegel, 1979). The purpose of the experiments reported in this chapter determine whether the notion of a spatial mental model is a useful construct in predicting and explaining the retention of spatial information over very long retention intervals. Ecological validity is also an important reason for looking at the very long term retention of spatial information. Many areas of cognitive psychology rely on studies of long-term memory over very short retention intervals, minutes or sometimes hours. In real life the execution of plans, for navigation and many other activities require the retention of spatial information over periods of days, weeks and years. So it is important that findings from short retention intervals can be extended or mapped onto evidence from studies using very long retention intervals.

Bahrick (1984) showed that deeply learned knowledge could remain stable over very long periods of time. His interpretation of this finding in terms of a 'permastore' has been challenged by Ulric Neisser. Neisser interpreted the same findings in terms of the retention of a schematic knowledge which permits specific details to be reconstructed. Studies by Cohen, Conway and Stanhope (1992) in several different areas of formal education have confirmed that high-level schematic knowledge is well-preserved. However, the same studies have also shown



that certain kinds of specific information, such as names, can be well remembered; a finding which is difficult to account for in terms of schema reconstruction.

## 5.1 Very long term memory for pictures and places

Several early studies of very long-term retention of spatial information explored recognition memory for pictures. Shepard (1967) presented a number of subjects with up to 600 pictures. With a mean duration of presentation of 5.9 seconds recognition memory was very high. On a forced choice recognition test median accuracy was 98.5% on an immediate test, 100% after two hours, 93% after 3 days, 92% after 7 days and 57% after 120 days. While experiments like those of Shepard show that visual memory for a picture is excellent they say very little about memory for the spatial or structural characteristics of the pictures.<sup>1</sup>

Virtually all other studies of very long term retention of spatial information have focused on memory for places acquired from navigation (Bahrick, 1983; Chase, 1983; Herman et al., 1979; Thorndyke & Hayes-Roth, 1982). The most comprehensive study of this kind by Harry Bahrick (1983) looked at 851 students' memories of Delaware up to fifty years after leaving the city. Bahrick found that most spatial knowledge was learned in the first three weeks and that when other factors (such as access to a car) were taken into account there were no reliable gender differences in spatial learning or memory. Knowledge of landmarks and buildings was gained more quickly than for the street network. This probably reflects the needs of the student population and the unsystematic arrangement of Delaware street names.<sup>2</sup> The rate of forgetting was related to the rate of acquisition with street names and street order forgotten faster than landmark or building knowledge. A more recent experiment looked at memory for an unfamiliar outdoor environment learned in one or two short trips over a period of five months (Herman et al., 1987). Taking a second trip through the environment did not improve bearing or distance estimates in the environment. Bearing and distance estimates of eight and eleven year olds declined from the spring to autumn retention interval, while those of nineteen year olds did not. These results were interpreted in terms of older subjects possessing more advanced and more flexible 'co-ordinate frames'. In other words retention is preserved through better integration of spatial information during encoding. An alternative explanation, that adults were better at using landmark pictures as retrieval cues was also considered.

---

<sup>1</sup> Nearly all of Shepard's materials depicted a single clear object. This, coupled with the short presentation times, might suggest that it is primarily the visual rather than the spatial or structural characteristics that are forming the basis of recognition.

<sup>2</sup> Bahrick cites research by Devlin (1976) where people show superior memory for street names and locations. Devlin's study was carried out in the town of Idaho Falls where streets are organized alphabetically.



What changes are likely to occur in spatial representations over time? The integration of spatial information into a single, coherent structure appears to be an important factor in the retention of spatial information. Information which is better integrated should be better retained. For landmarks acquired in a particular sequence (e.g. from a description or from traversing an actual route) this may mean that early landmarks are better remembered than later ones. An integrated spatial representation may also preserve global spatial relations better than local ones (because global spatial relations can usually be derived from several different local spatial relations).

## 5.2 Recall, reconstruction or guesswork

A major methodological problem in studies of long-term memory for meaningful information has already been hinted at. Recall may be strongly influenced by long-term memory schemas. Schematic knowledge may have driven the original learning episode or it may be used to reconstruct details. Generate-recognize theories of recall (Jones, 1978; Watkins & Gardiner, 1979) suggest a mechanism for this recall process (see section 1.6.2). Highly typical items are likely to be remembered better because they can be easily generated from schematic knowledge. Once generated they may be recognized or rejected. Indirect recall routes, such as generate-recognize, are also more likely to be employed when direct recall routes (Jones, 1978; 1982; 1987) are less accessible (e.g. at very long retention intervals). A different, but related problem is the possible use of 'guesswork' strategies by subjects. The use of indirect (generate-recognize) recall routes may be difficult to distinguish from guesswork, for instance when the to-be-remembered material contains a high proportion of typical items.

The studies reported below attempt to account for schema-based recall and the use of guesswork strategies where subjects may mistake typicality or plausibility for genuine recognition. In all three experiments reported below errors of commission and errors of omission are reported separately. Guesswork might be indicated by errors of commission for typical items. Whereas schema-based recall might be indicated by errors of omission for low typicality items. In addition, for Experiment 7 probability of landmark recall is compared with typicality and frequency ratings carried out for the study. In Experiment 8 performance of experimental subjects was compared with that of a second group of subjects who had not taken part in the original learning episode. This second group of 'baseline' subjects were given explicit guesswork instructions (i.e. told to draw or generate plausible or typical items on each experimental task). They thus acted as a baseline for the level of 'recall' which could be predicted by guesswork strategies.



### 5.3 Very long term memory for spatial mental models

Common findings from studies of long term retention in spatial and non-spatial domains suggest that the ability to integrate information into structured, stable knowledge structures is very important. An experiment by Waddell and Rogoff (1987) provides further support for this view. They found that intentional strategies relating object information to spatial location improved memory for location information. However, the role of schematic information at retrieval cannot be ruled out as a factor in superior memory for organized information. A number of studies have shown that schemas can influence recall or recognition even when activated after encoding has taken place (Anderson & Pichert, 1978; Pichert & Anderson, 1977).

If the construction or elaboration of a spatial mental model are significant processes in the integration of knowledge structures during learning there may be evidence of these processes even at very long retention intervals. The spatial mental model account would propose that the order of spatial mental model construction would reflect the degree of integration and elaboration of elements in a given spatial mental model (refer to the discussion of these issues in Chapter 2 and the experiments reported in Chapters 3 and 4). The spatial mental model should preserve structural or spatial information important for the original learning episode. Information incidental to the requirements of the original learning episode (e.g. about the language of a description where the original learning episode requires the manipulation or retention of spatial information) should be worse preserved than information which is integrated into a spatial mental model (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). This is because the construction and elaboration of a mental model (spatial or otherwise) requires both time and effort (de Vooght & Vandierendonck, 1993; Gernsbacher, 1990; Gernsbacher, 1991; Tardieu et al., 1992; Wilson et al., 1993).



## 5.4 Experiment 7

This experiment is a long-term follow-up of memory for spatial information originally learned in the form of short narratives. In the original study (Experiment 1, presented in Chapter 3) subjects were presented with short stories describing a village called 'Lower Barking'. The mean retention interval between the original learning episode and the subsequent follow-up was just over seven months. In the original experiment half the subjects received scrambled stories, while the remaining half received intact, unscrambled stories. Of the unscrambled stories one half were written from a route perspective and the other half took a survey perspective (see Taylor & Tversky, 1992a; Thorndyke & Hayes-Roth, 1982). In this follow-up only subjects who received unscrambled stories were tested. In the original experiment recall of landmarks in a drawing task was at ceiling for subjects in the unscrambled conditions but not for the subjects in the scrambled conditions. For this reason only the unscrambled conditions were re-tested for this experiment.

In this study subjects from the original experiment were tested in two ways. Firstly, they were asked to recall information from the original study episode by drawing the learned environment from memory. Secondly, they were asked to perform a recognition test for the original sentences in the story they had read. The recognition test consisted of eighteen sentences in a random order; six verbatim sentences from the original text, six paraphrases of sentences from the original text and six inference sentences. The inference sentences contained spatial relations that were not explicit in the original texts but which were accurate with respect to the spatial layout of the village they described. All eighteen sentences took the same perspective (route or survey) as the text read by a given subject in the original experiment.

The spatial mental model account outlined in Chapter 2 proposes that recall from long-term memory should be related to the predicted order of spatial mental model construction. Two measures of such an order effect are possible; the drawing orders of items where recall is close to ceiling or the probability of recalling an item where memory is below ceiling. A second prediction from the spatial mental model account is that (for determinate descriptions) memory for the language of a description will be poorer than memory for the spatial relations. This is because people will tend to abandon any effort to remember the language of a description if they are able to construct a single, coherent model of the situation (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). Results from Experiment 6 (see Chapter 4) suggest that there are also situations where memory for an indeterminate description is not based on a representation of the language of a description. Instead memory for indeterminate descriptions may rely primarily on memory for the construction of a spatial mental model or even for the structure of a spatial mental model. A final prediction is that long-term memory for a spatial mental model should not differ between route and survey versions of the text. A spatial mental model is a perspective free representation (Taylor & Tversky, 1992a; Tversky, 1991), though it is possible to generate



perspective from a spatial mental model. By contrast an image or possibly a representation of the language of the text may preserve information from the original text perspective.

### 5.4.1 Method

#### Subjects

Sixteen subjects (eleven male and five female) took part in the experiment. All sixteen subjects had taken part in a previous experiment requiring them to read and remember a short narrative describing a village (see Experiment 1 in Chapter 3). The mean retention interval between the original learning episode and the subsequent follow-up was 220 days (the standard deviation was 25 days). Nine of the subjects had been given the survey perspective version of the narrative presented in Experiment 1 and seven had been presented with route perspective version. All seventeen subjects had been presented with unscrambled versions of the original texts (i.e. the narratives had been presented in the original or natural order). An additional ten subjects provided ratings for the eight landmarks used in this experiment.

#### Materials

Subjects were provided with a blank sheet of paper in the first test stage. In the second, final test stage each subject was presented with a recognition test (note that none of the subjects received a recognition test in Experiment 1). The recognition test materials consisted of eighteen sentences describing the village mentioned in the original stories that subjects had read (see Experiment 1, Chapter 3). Two sets of recognition materials were prepared, one for each of the two perspective versions (survey or route) used in the original texts. Each set of recognition materials contained six verbatim sentences, six paraphrase sentences and six inference sentences. The verbatim sentences were taken at random from the original route or survey version of the narrative that subjects had read. The paraphrase sentences were based on six sentences chosen at random from the original narrative. The paraphrases incorporated word or phrase substitutions that preserved the meanings of the original sentences. The inference sentences described spatial relationships between landmarks in the village which were not explicitly mentioned in the original narrative. Thus all of the eighteen sentences described genuine spatial relationships between landmarks in the village. For the survey perspective recognition materials all eighteen sentences took a survey perspective. For the route perspective recognition materials all eighteen sentences took a route perspective. The eighteen sentences were presented in a random order on a single sheet of paper. An example of each type of sentence for both the survey and route text is shown in Figure 5.1.



Verbatim sentence (survey text).

On the northern corner of the main road and Dog Kennel Lane stood the village Post Office.

Paraphrase sentence (survey text).

Dog Kennel Lane ran all the way round the Common in the centre of the village.

Inference sentence (survey text).

The Deserted Farmhouse was Northwest of St. Malcolm's Church.

Verbatim sentence (route text).

To Eric's left, on the corner of the main road and Dog Kennel Lane, stood the village Post Office.

Paraphrase sentence (route text).

Eric slowed down as he drove into the village from the West, here the main road had made a T-junction with Dog Kennel Lane.

Inference sentence (route text).

Eric drove past the Post Office as he went round Dog Kennel Lane.

Figure 5.1: Examples of the recognition test materials for route and survey text.

In addition to these materials ratings were obtained for the four buildings and the four roads in the original texts. These ratings were obtained from a sample of nine people who were not being tested for long-term retention. They were asked to rate the eight landmarks according to how 'typical' and how 'common' they were for a traditional English village.

### Procedure

There were two stages to the experiment. Before beginning the first stage subjects were asked to think back to the original experiment they had taken part in. They were reminded that it had involved reading a short narrative describing a village called 'Lower Barking'. The first test stage involved attempting to draw the village from memory. Subjects were asked to draw as many roads and buildings as they could remember and, where possible, to label landmarks as they drew them. In stage two subjects were asked to perform a recognition test. Each subject was given an appropriate set of recognition materials (sentences with either a route or survey perspective). For each sentence they were asked to circle the response 'yes' or 'no' according to whether they recognized the *exact* wording of the sentence and not simply if they thought the sentence was true. Prior to taking part in the experiment ten of the sixteen subjects had also completed a Vividness of Visual Imagery Questionnaire or VVIQ (Marks, 1973).



5.4.2 Results

Scoring

Scoring free recall in a drawing task is difficult. As in previous experiments (see Chapter 3) subjects were instructed to label roads or buildings as they drew them. Unfortunately, and in contrast to the results of previous experiments reported here, most subjects were unable to label many of the features or landmarks that they drew, and it was not possible to identify the majority of the unlabelled landmarks. Memory for road names was particularly poor. For this reason it was not possible to analyze drawing orders. Nor was it possible to analyze spatial location because the majority of roads (which connect the other landmarks) were not labelled. Instead the initial analysis was confined to landmarks and features that subjects were able to label. Two types of scoring were employed. The first was exact or verbatim recall of a landmark that was drawn. Exact landmark recall was scored if the original landmark name was reproduced verbatim (e.g. 'Rose and Crown Inn' or 'Dog Kennel Lane'). The second scoring procedure was for partial recall. Partial recall was scored if part of the exact name or a closely associated name was produced (e.g. 'Pub' for 'Rose and Crown Inn', 'Greyhound Lane' for 'Dog Kennel Lane').

Recalling landmark names

Figure 5.2 shows recall of landmark names in subjects drawings according to how they were classified (exact recall, partial recall, errors of commission or unclassified):

	Number of landmarks drawn	Exact recall	Partial recall	Errors of Commission	Unclassified
Recall	8.75	8.7%	28.2%	13.4%	49.7%

Figure 5.2: Mean number of landmarks drawn and the proportion classified as exact recall, partial recall or errors of commission.

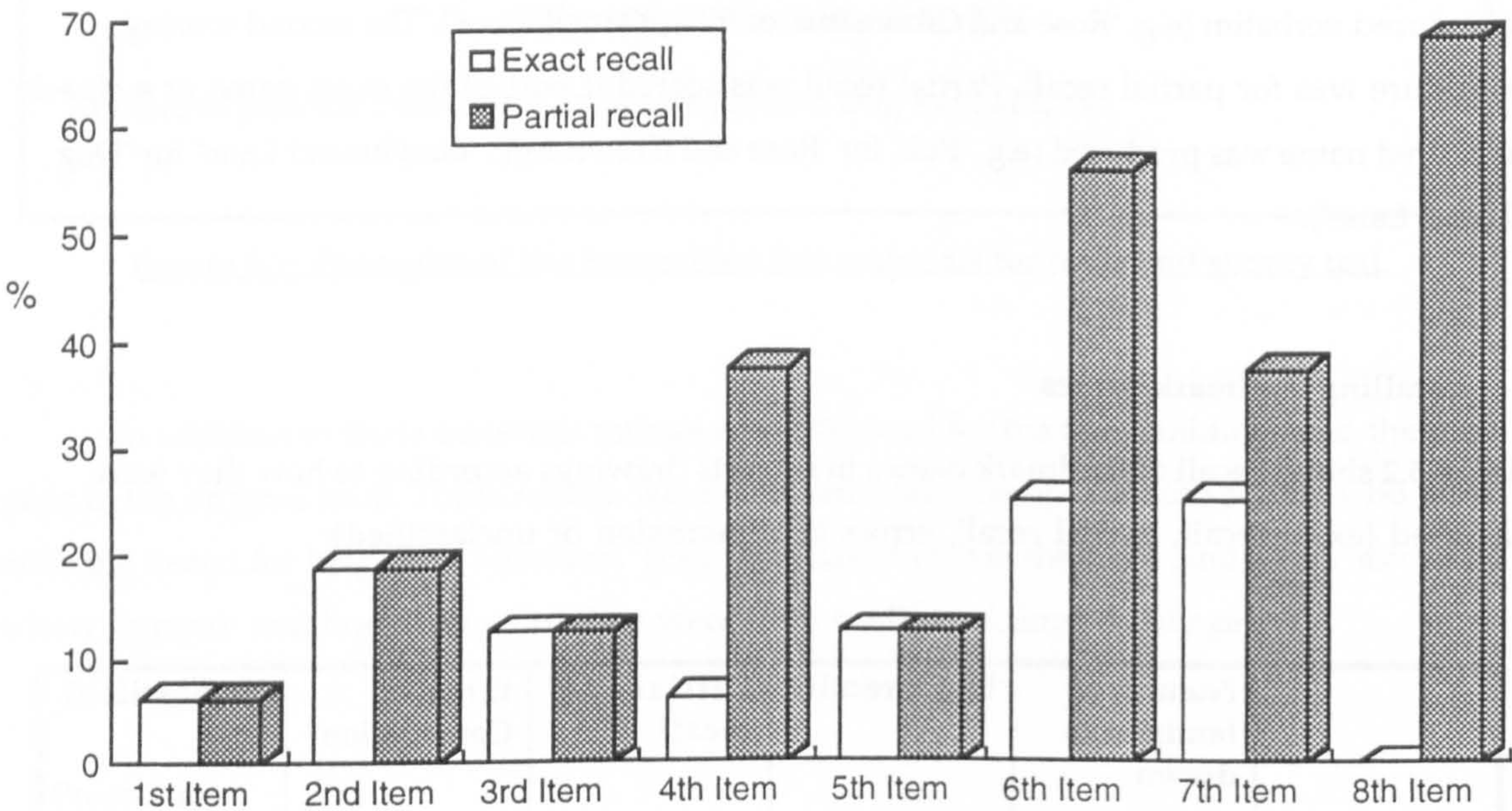
Overall levels of recall for the landmarks is difficult to judge accurately. Nearly half the landmarks drawn were not labelled and only in a small number of cases were landmark names recalled accurately. Partial recall however, is over twice as high as the proportion of errors of commission (this is significant using the Wilcoxon signed-ranks test;  $N = 16$ ,  $Z = 2.95$ , one-tailed  $p = 0.0016$ ). This suggests that in most cases subjects are able to discriminate between plausible landmarks and landmarks that were actually present.



Subjects who had read route versions of the text did not differ significantly from subjects who read survey versions in their memory for landmarks. This was true for the proportion of exactly recalled names (Mann-Whitney  $U = 28.5$ ,  $N = 16$ ,  $Z = 0.34$ ; two-tailed  $p = 0.73$ ), partially recalled names ( $U = 25.5$ ,  $N = 16$ ,  $Z = 0.66$ ; two-tailed  $p = 0.51$ ), errors of commission ( $U = 25.5$ ,  $N = 16$ ,  $Z = 0.86$ ; two-tailed  $p = 0.39$ ) and unclassified landmarks ( $U = 24$ ,  $N = 16$ ,  $Z = 0.83$ ; two-tailed  $p = 0.41$ ). There is no evidence of any difference in accuracy between subjects who read the route text and those who read the survey text version.

### Recalling landmark names

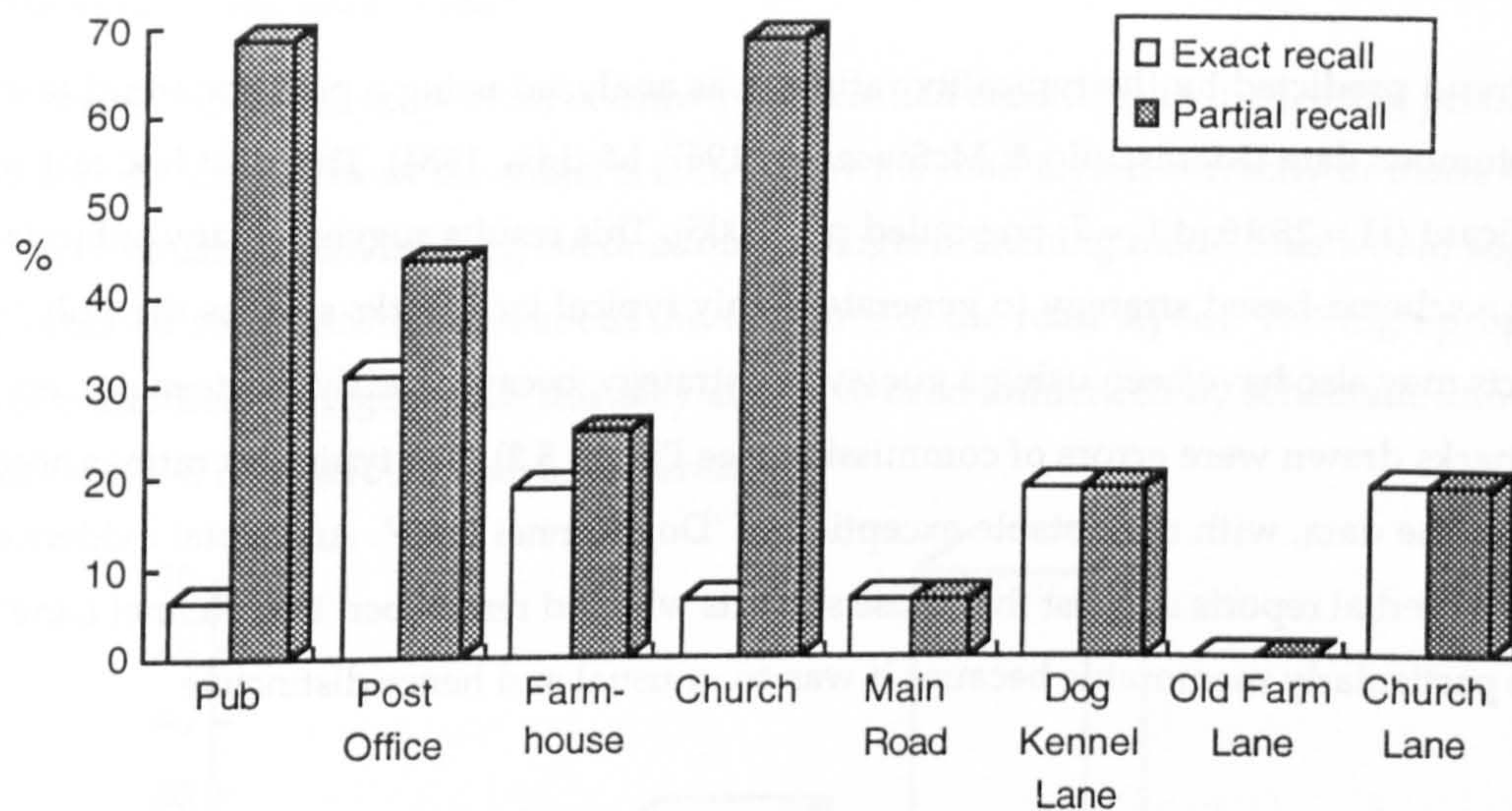
Figure 5.3 shows the probability that the name of a landmark was recalled plotted against the order items were presented in the original unscrambled survey or route text. These figures are given separately for exact recall and partial recall. The spatial mental model account predicts that items mentioned earlier in the original text will be remembered best.



**Figure 5.3** The probability that exact or partial landmark names are recalled against serial position of the landmark in the original text.

Exact landmark name recall is very poor. Nor is there any evidence of any trend relating serial position of landmarks in the text and probability of recalling the exact name. Partial recall of landmark names is significantly greater. There is no evidence of a trend between probability of recall and the predicted order of spatial mental model construction. In fact, partial recall of landmark names appears to be negatively related to the predicted order of spatial mental model construction. However, it is possible this may be the result of a coincidental relationship between the typicality of a landmark and their position in the text. Figure 5.4 shows the probability of recalling each landmark in the original text independently of serial position in the text.





**Figure 5.4 : The probability of exact or partial recall by type of landmark.**

From Figure 5.4 it appears that highly typical or schema consistent landmarks such as the pub, post office and church showed higher levels of recall and in particular partial recall. To test this the probability of partial recall was tested against two ratings; how typical and how common a given landmark was for a traditional English village. Ratings for typicality and commonness both predicted very similar orders of recall (see Figure 5.5).<sup>3</sup>

Landmark:	How typical (mean rating):	How common (mean rating):
<i>Pub</i>	1.11	1.33
<i>Church</i>	1.11	1.88
<i>Post Office</i>	1.67	2.11
<i>Farm House</i>	2.67	3.00
<i>Old Farm Lane</i>	3.00	3.77
<i>Church Lane</i>	3.11	2.89
<i>Main Road</i>	4.22	5.00
<i>Dog Kennel Lane</i>	5.67	6.56

**Figure 5.5: Mean ratings for typicality and commonness of the eight landmarks**

<sup>3</sup> Ratings made for the buildings used generic category names (pub, church and so on) because this was the criterion of successful partial recall. Rating for the roads did not because it is not clear what generic category a road name like Dog Kennel Lane or Old Farm Lane could be classified under.

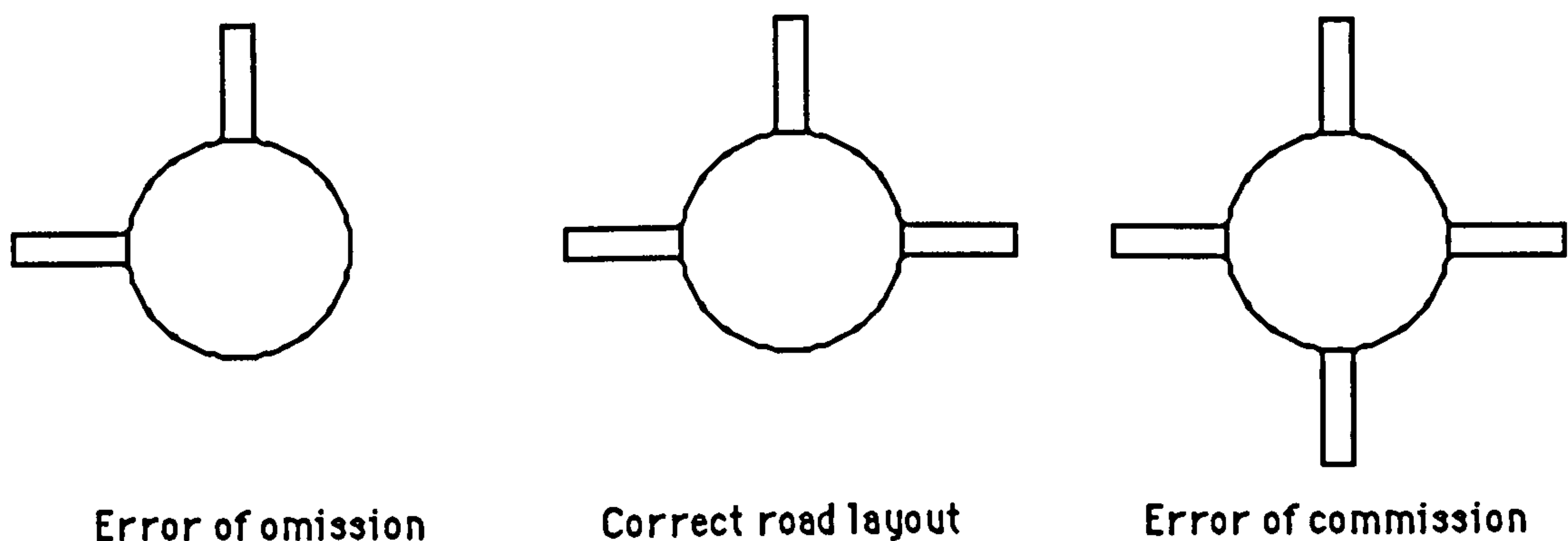


This trend predicted by the typicality ratings was analyzed using a post hoc trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). This post hoc test was significant ( $H = 28.16$ , d.f. = 7; one-tailed  $p < 0.005$ ). This results suggests many subjects were using a schema-based strategy to generate highly typical landmarks such as the pub. Some subjects may also have been using a guesswork strategy, because nearly fourteen percent of landmarks drawn were errors of commission (see Figure 5.2). The typicality ratings account very well for the data, with the notable exception of 'Dog Kennel Lane'. Anecdotal evidence from subjects' verbal reports suggest that those subjects who did remember 'Dog Kennel Lane' found the name particularly memorable because it was so unusual and hence distinctive.

## Recalling the road layout

### Scoring

The two scoring procedures used above take into account only part of the information recalled by subjects due to the large proportion of landmarks left unclassified. The majority of these unclassified landmarks were unlabelled roads drawn by subjects (and so are not included in Figure 5.4). Both the route and survey versions of the original story described a circular road ('Dog Kennel Lane') which surrounded the village common. Radiating from the central road were three other roads (the 'Main Road', 'Old Farm Lane', and 'Church Lane'). This distinctive road layout suggested a third scoring procedure. Each of the free recall maps drawn by the subjects was classified as correct (if three roads radiating from a central, circular road were drawn) or incorrect (if any other arrangement of roads was drawn). In addition, the incorrect layouts were classified as errors of omission (if one or more roads was missing), errors of commission (if additional roads were added) or other (if the road layout did not incorporate a circular central road with other roads branching off it). The scoring procedure is illustrated in Figure 5.6.

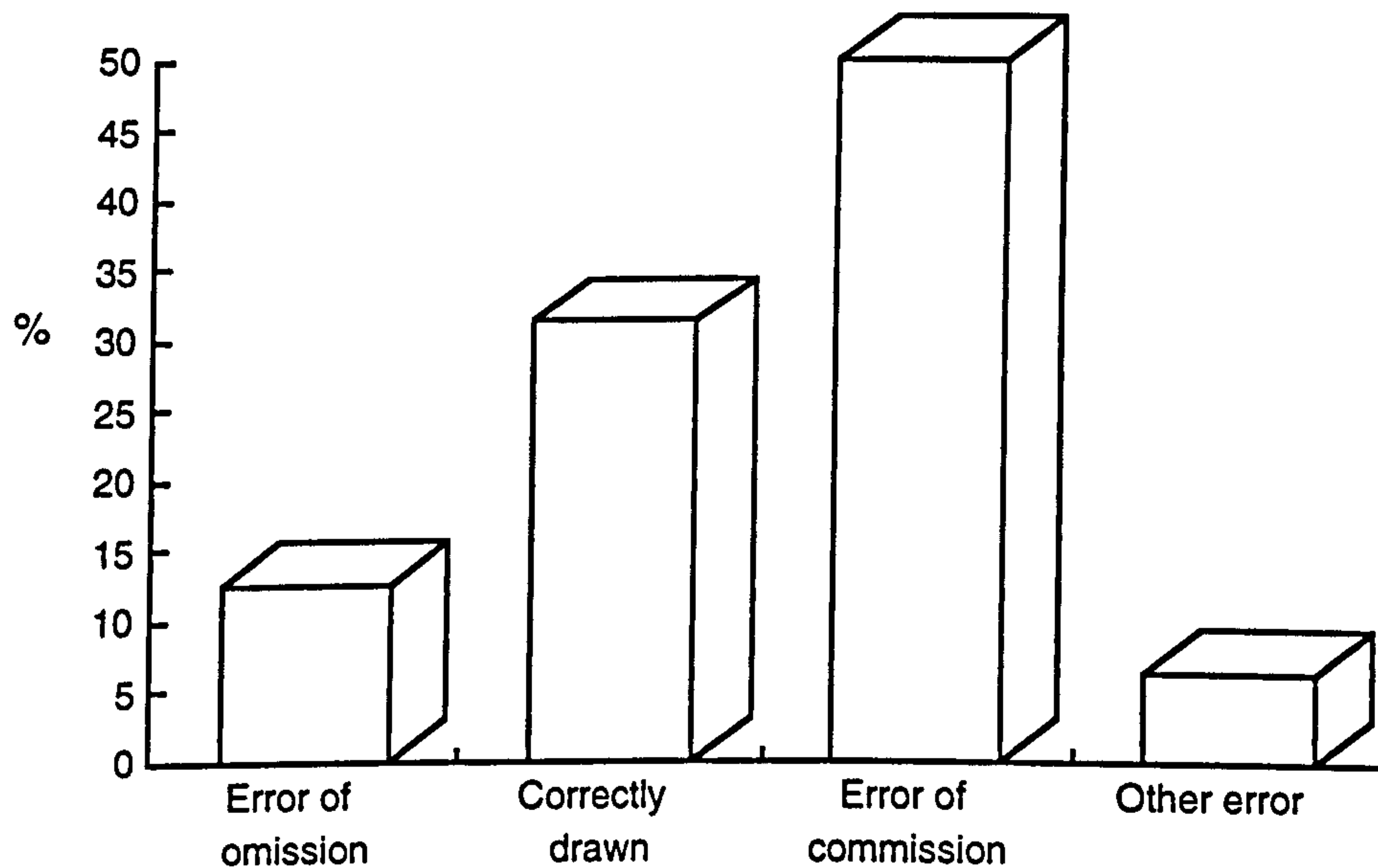


**Figure 5.6: Examples of road layouts drawn by subjects.**



### *Memory for the road layout*

Figure 5.7 below shows the proportion of drawings classified according to the scoring procedure described above. All but one of the subjects either drew the road layout correctly or made errors of commission or omission (i.e., missing out or adding a single branching road). This would suggest that they retained some information about the structure of the road layout. The high proportion of errors of commission suggests that this may also have been influenced by schematic knowledge (e.g. about common road layouts such as crossroads).



**Figure 5.7: Percentage correctly recalled road layouts, errors of omission, errors of commission and other (unclassified) errors.**

Route and survey perspective conditions did not differ greatly in their memory for the road layout. This was confirmed with a Chi Square test on the drawing classifications shown above between the survey and route conditions (Chi Square = 1.67, d.f. = 2, two-tailed  $p = 0.43$ ).

### **Recognition memory for the language of the text**

#### *Scoring*

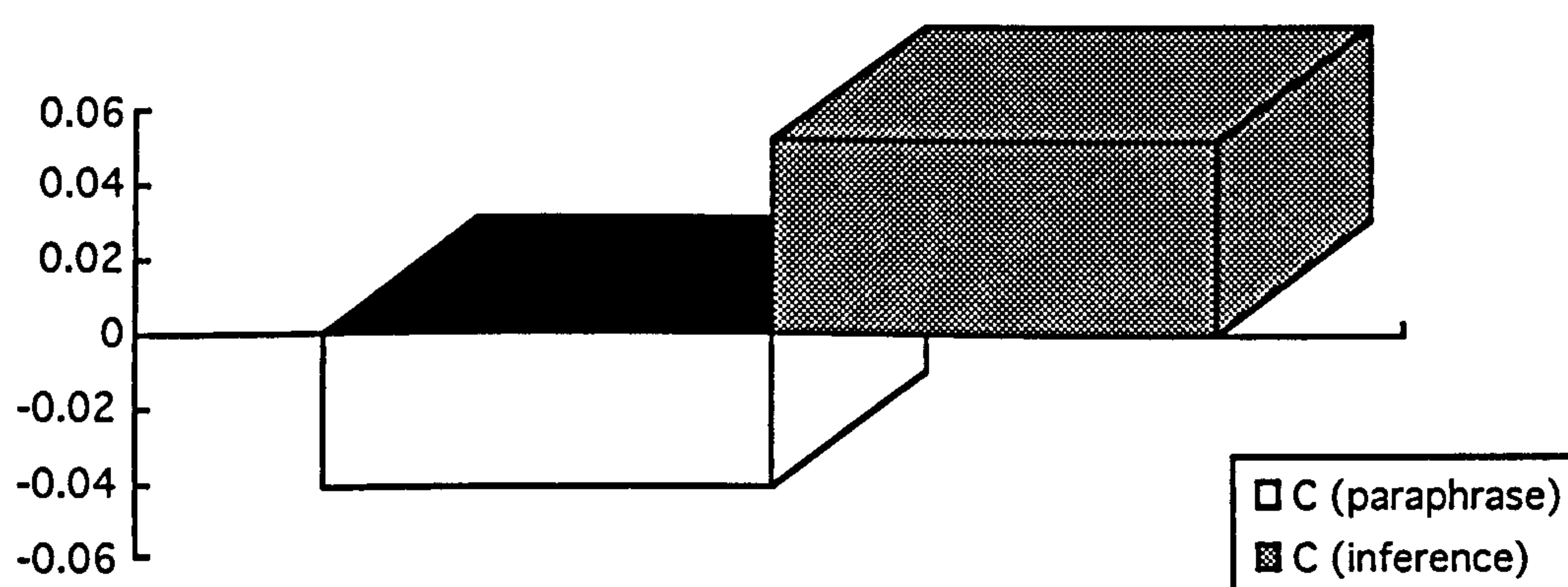
Recognition scores were calculated out of six for each sentence type (verbatim, paraphrase or inference). Correct recognition was scored if a subject responded 'yes' to a verbatim sentence or 'no' to a paraphrase or inference sentence (subjects were instructed to respond 'yes' only to sentences they recognized as having exactly the same wording as the original story they had read). However, it was not possible to determine recognition accuracy from raw scores (because of probable biases to 'yes' or 'no' responses). For this reason a measure of signal detection was adopted. In order to avoid making strong assumptions about the distribution of responses in this sample a non-parametric measure of signal discriminability was employed (Altham, 1973; Hammerton & Altham, 1971). This measure 'C' is monotonically related to  $d'$ . Where



discriminability is perfect  $C = 1$ , where responses are random  $C = 0$  and for consistent error  $C = -1$ . Using raw recognition scores it was possible to generate discriminability scores for each subject between verbatim and paraphrase and verbatim and inference sentences. These are referred to as  $C$  (paraphrase) and  $C$  (inference) respectively.

### *Discriminating verbatim sentences from paraphrase and inference sentences*

As outlined in the scoring procedure described above measures of recognition accuracy were calculated for each subject. These resulted in  $C$  (paraphrase) and  $C$  (inference) values which show the level of discriminability between verbatim sentences and paraphrase or inference sentences respectively. Figure 5.8 shows mean  $C$  (paraphrase) and mean  $C$  (inference) for the recognition scores produced by subjects.



**Figure 5.8 Mean discriminability ( $C$ ) in recognition performance between verbatim sentences and paraphrase or inference sentences.**

$C$  scores for the recognition test are very close to zero, suggesting that subjects are not able to discriminate between verbatim sentences and paraphrase or inference sentences (in fact there is a slight tendency for subjects to prefer paraphrases over verbatim sentences). The possibility that discriminability between verbatim and inference sentences was above chance was tested using the binomial test (Siegel & Castellan, 1988). This test assumes that under the null hypothesis discriminability is equally likely to be above or below zero. This test does not take account of the magnitude of these differences, however, because the overall mean is so close to zero it is extremely unlikely that these difference would have a significant effect. Discriminability was not significantly greater than zero for  $C$  (inference) scores ( $N = 16$ , one-tailed  $p = 0.61$ ). Therefore there is no evidence that subjects were able to discriminate between verbatim sentences and paraphrase or inference sentences on the recognition test.

There were no significant differences on recognition between subjects who had route and survey versions of the text for  $C$  (paraphrase) scores (Mann Whitney  $U = 28$ ,  $N = 16$ ,  $Z = 0.39$ , two-tailed  $p = 0.70$ ) or  $C$  (inference) scores ( $U = 29$ ,  $N = 16$ ,  $Z = 0.27$ , two-tailed  $p = 0.79$ ). Thus



text perspective did not appear to influence subjects ability to discriminate verbatim sentences from either paraphrase or inference sentences.

### Correlations

Neither retention interval nor age of subject correlated with the measures of landmark recall or recognition discriminability used above. Vividness of visual imagery (VVIQ scores) did not correlate significantly with exact recall of landmark names (Spearman's  $Rho = 0.349$ ,  $N = 10$ , two-tailed  $p = 0.324$ ), partial recall ( $Rho = 0.255$ ,  $N = 16$ , two-tailed  $p = 0.47$ ) or errors of commission ( $Rho = 0.441$ ,  $N = 16$ , two-tailed  $p = 0.21$ ). Note that high vividness is associated with low VVIQ scores. Similarly low correlations are obtained between VVIQ scores and discriminability between verbatim and paraphrase ( $Rho = -0.294$ ,  $N = 16$ , two-tailed  $p = 0.41$ ) or verbatim and inference sentences ( $Rho = -0.427$ ,  $N = 16$ , two-tailed  $p = 0.23$ ). There is no evidence that vivid imagers remember more landmark names (in fact high vividness scores are weakly associated with poor landmark recall). Nor is there a significant relation between sentence discriminability and vivid imagery.

### 5.4.3 Discussion

The spatial mental model account generated three predictions about the long-term retention of information derived from texts. The first prediction was that probability of recalling landmarks in a drawing task should reflect the order of spatial mental model construction. However, the probability that landmarks (in this case roads and buildings) were recalled in the drawings made by subjects was not related to the spatial mental model construction order. Instead ratings of how typical or common the items were shown to be strong predictors of the probability that a given landmark was recalled. High typicality or schema consistent landmarks (in the case on an English village a pub or a church) were many times more likely to be recalled than low typicality items such as 'Old Farm Lane'. These findings provide support for the involvement of schematic knowledge in reconstructing information during recall (e.g. Bartlett, 1932; Anderson and Pichert, 1978; Brewer and Treyens, 1981). The low proportion of errors of commission suggest that the mechanism for this reconstruction process may be a generate-recognition recall procedure (Jones, 1978; Jones, 1982; Watkins & Gardiner, 1979). People use schematic knowledge to generate possible landmarks, but are nevertheless able to reject plausible but inaccurate alternatives such as 'school' or 'garage'. It appears that the influence of schematic knowledge on the probability of recall makes it very difficult to detect order effects on the probability of recall. It also suggests that the drawing and recall order effects in Experiment 1 (see Chapter 3) cannot be explained in terms of the typicality or perceived frequency of the landmarks themselves. Thus there is no direct evidence that people are influenced by the order of spatial mental model construction. Subjects made relatively few errors of commission, though, suggesting that some memory for the original learning episode remained.



The second prediction was that memory for the language of the text should be less well retained than memory for a spatial mental model. Partial support for this prediction is obtained. Discrimination between verbatim and both paraphrase and inference sentences is at chance on the recognition test. However, to confirm this prediction would require direct evidence that what subjects do recall is based on memory for a spatial mental model.

The third prediction was that the perspective taken by the original text should not influence memory for spatial information. Text perspective had no effect on any of the recall or recognition scores used in this study. It is possible that differences may have been obscured by the small sample sizes or the poor recall or recognition performance of subjects. There is certainly no evidence of any dramatic differences in recall or recognition between subjects. Similarly, there appears to be no influence of age of subject, retention interval or vividness of visual imagery as rated by the VVIQ (Marks, 1973). While these factors may influence spatial memory or memory for landmark names these results suggest that their influence is relatively small.

Poor recall for landmark names in the drawing task (almost half of the features drawn were unlabelled) may have resulted in recall of spatial information being underestimated. A different scoring procedure - memory for the distinctive road network - showed that nearly every subjects recalled the road network (provided errors of omission or commission of a single road are included). Determining whether this level of recall is possible by chance is difficult. Memory for the road layout would be consistent with memory for an image or for a spatial mental model. However, the large proportion of errors of commission (in this case the addition of a single road) would suggest a structural or spatial representation rather than a visual one.

This experiment provides mixed evidence for the predictions based on the spatial mental models account. Recognition memory for the language of the text is very poor. Text perspective does not influence recall or recognition. There is no support though for the most important prediction that recall on the drawing task would be influenced by the order of spatial mental model construction. Whether a landmark is consistent with schematic knowledge about the environment being drawn is a far better predictor of recall. There are very few errors of commission in recalling landmark names suggesting that subjects maybe adopting a generate-recognition recall route (Jones, 1978; Jones, 1982; Watkins & Gardiner, 1979). In addition, memory for landmark names and for the road network suggests that recall may be based on a spatial representation derived from the original text. This representation may be in the form of a spatial mental model or possibly an image. Future experiments will have to take greater account of the role of schematic knowledge on reconstruction or recall in order to determine any direct evidence that subjects are able to remember a spatial mental model at retention intervals of several months.



## 5.5 Experiment 8

Experiment 7 failed to produce direct support for the proposal that spatial mental models influence spatial memory at very long retention intervals. In this experiment a number of different tests are employed in order to avoid some of the methodological problems associated with studying memory at very long retention intervals. This was accomplished by following up subjects who had taken part in an experiment requiring them to learn a map of a university campus and a route taken across the same map. Subjects were tested between thirteen and nineteen months after they had originally learned the map and the route. The original experiment investigated differences in people's ability to integrate route information with survey information derived from the map. The route was presented either in a *verbal* form (a sequence of sentences) or in a *visual* form (a sequence of landmarks). In the original study mode of presentation of the route had little or no effect on recall. The only significant difference between the two conditions was that people who had received the verbal version of the route were quicker to answer written questions about the spatial information they had learned (Baguley, 1992).

### 5.5.1 A brief description of the original task

The original task was divided into five stages. In the first stage of the experiment subjects were asked to study a map of an imaginary university campus. Each subject was given four minutes to study the map. In the second stage the subjects were asked to study a route taken across the same imaginary university campus. Subjects were randomly allocated to one of two groups for the experiment. Both groups were presented with an identical map in stage one. In stage two, however, the route was presented in a different format for each group. The first group received the route written as a short descriptive text. The second group received the route in a visual form (as a series of landmarks connected by arrows). In both conditions subjects were given three minutes to learn the route. The map learning and route learning stages were followed by three test stages which were identical for the two conditions. The two conditions differed only in the presentation of the route in a *verbal* or a *visual* form.

In stage three subjects were asked to list, in order, the named features on the route. In stage four subjects were given a questionnaire. The questionnaire consisted of fifteen multiple choice questions involving information about both the original map and the route. In the fifth and final test stage subjects were given a copy of the original map and asked to draw the route upon it. In addition subjects were asked to describe what they were thinking as they marked the route on the map and a verbal protocol was recorded.



Subjects receiving the route in a written form were significantly quicker in answering the questionnaire in the second test stage. One possible explanation of this effect was that these subjects gained a reading time advantage from the written questionnaire, through already having formed a textbase of the route (van Dijk & Kintsch, 1983). There were no other significant differences between the two conditions.

### 5.5.2 The experiment

The purpose of the study was two-fold. Firstly to establish whether it is possible to eliminate or otherwise take account of the influence of schematic knowledge on recall. Secondly, to discover whether recall is influenced by the predicted order of spatial mental model construction.

One problem with testing the long-term retention of meaningful spatial information (e.g. a map rather than an arbitrary spatial array) is to determine what level of performance is possible by chance or through guesswork. People may consciously decide to guess or their memory for a real university campus might influence the memory for the experimental situation just as knowledge about typical landmarks influenced recall for the village in Experiment 7. To help eliminate this possibility performance of subjects in this study was compared with the performance of subjects who had not learned the original map or route. These subjects were to provide a 'baseline' of performance. The baseline subjects were given 'guesswork' instructions in place of instructions to remember the original learning experience. The guesswork instructions were for subjects to complete the task with the most plausible or likely responses.

Four different tasks were adopted in order to minimize the involvement of schematic knowledge by testing memory in different ways. Generate-recognition recall routes (Jones, 1978; Jones, 1982; Watkins & Gardiner, 1979) probably inflate recall for schema consistent landmarks or features. Providing cues in some tests or assessing recognition memory on others should reduce the impact of generate-recognition routes on recall. In addition performance of subjects in the baseline condition should provide an indicator of the level of recall or recognition performance possible by chance. The first task was a repetition of the free recall drawing task in Experiment 7. The second task assessed recognition memory for fifteen of the landmarks on the original map. The third task assessed cued recall for the location of landmarks from the original university campus map. This was done by providing subjects with landmark names and asking them to place these name labels in the appropriate location on an unlabelled map of the university campus. The fourth and final task was for subjects to draw the route they had learned onto an intact copy of the original map. Only in the first task are subjects required to generate landmark names. Hence for the remaining three tasks *intrinsic* cues (direct cues relating to information encoded in



the original study episode) are available. This should reduce the impact of *extrinsic* cues (not encoded in the original study episode) such as those generated using schematic or background knowledge (Jones, 1978; 1982; 1987).

Memory for a spatial mental model should be reflected in several ways. Landmark names and locations should be recalled and recognized above chance (as determined by performance in the baseline condition). This is necessary, but not sufficient to demonstrate memory for a spatial mental model. Memory for the route should also reflect the predicted order of spatial mental model construction. Landmarks entered into a spatial mental model early on should be better recalled than those integrated into the representation at a later stage. As in previous experiments the order of spatial mental model construction may also influence drawing orders adopted by subjects (Taylor & Tversky, 1992a). Finally, the recall of the route should not be influenced by the mode of route learning. If *verbal* and *visual* conditions differ in the extent to which the route is recalled these differences could be explained in terms of memory for the visual aspects of the picture or for the language of the verbal. If route recall for both the verbal and visual conditions show the trend predicted by the order of spatial mental model construction this would support a common amodal, relatively abstract form of representation such as a mental model.

### 5.5.3 Method

#### Subjects

The twenty-five subjects for the experiment consisted of volunteers of varied age and background (most were graduate students or staff from the Open University). Seventeen subjects (seven male and ten female) had participated in a map learning experiment between thirteen and nineteen months previously. The remaining eight subjects (five male and three female) had not taken part in the original experiment and were unfamiliar with the experimental materials prior to the present experiment.

#### Design

The experiment investigated the long term retention of spatial information. Between subject factors were whether subjects had learned the map and route in the original experiment (experimental or baseline) and the original experimental route-learning condition (verbal or visual). Performance was compared over a number of measures of recall and recognition of spatial or spatially related information obtained from each of the three test stages.



## Materials

Subjects were tested using three sets of materials to be used in the second, third and fourth test stages, respectively. The first consisted of a list of forty-five feature or landmark names, of which fifteen were taken from the original campus map. The second was a copy of the original map with the names of all the buildings and roads deleted. To go with this map a set of thirteen labels was produced which could be placed on appropriate locations on the map. The final item was an intact copy of the original map.

## Procedure

Instructions for the experiment differed between experimental conditions. Subjects who had taken part in the original experiment were asked to remember the original study episode. They were reminded that the experiment had involved learning a map of an imaginary university campus and a route taken across that map by a character called Jerry. In the first test stage (free recall) they were asked to draw as much of the original map as they could from memory. In the second test stage (landmark recognition) they were given a list of forty-five features that might be found on a university campus. Subjects in the experimental condition were asked to circle those landmarks they recognized from the original experiment. In the third test stage (landmark location) they were given a copy of the map with the building and road names removed. They were then asked to place a set of thirteen labels according to their original locations on the campus. In the fourth and last test stage (route recall) subjects were asked to draw the route on a complete copy of the original map.

Instructions to the baseline subjects made no reference to the original experiment. Instead they were given 'guesswork' instructions. In stage one these subjects were asked to draw a typical example of a university campus. In stage two these subjects were asked to select the fifteen most likely features to be found on a map of a typical university campus. In stage three they were asked to place feature labels on the most appropriate or likely locations on the campus map. In the final stage the baseline subjects were asked to draw a likely or plausible route between any two campus locations of their choice. On those occasions where subjects did not indicate the direction of the route in their drawings this was recorded by the experimenter.

## 5.5.4 Results

Data from this study were analyzed in two ways. Firstly, data from the experimental condition were compared with those from the baseline condition. Data from the baseline condition are analyzed as if they had taken part in the original study. For example, responses are classified as incorrect or correct according to the same criteria used for those in the experimental condition even though there is no objective sense in which the baseline subjects can be said to have made an



error. Secondly, data from the experimental condition are analyzed separately to establish trends or patterns consistent or inconsistent with the predicted order of spatial mental model construction.

## Free recall

### *Scoring procedures*

All the maps obtained during free recall were analyzed using three sets of measures. The first measure was accuracy of object location. This was classified by dividing the original map into quadrants. Any labelled objects located on the map were then categorized as either correctly located (*correct*) or incorrectly located (*incorrect*). Some objects and features were unclassified (either because they were not labelled or because they were in two or more quadrants). Other features were not present on the original map and were classified as errors of commission (*commissions*).

The second measure was accuracy of recall for feature or landmark names. These were classified as either correct, partially correct, or errors of commission. A label was recorded as partially correct when it was semantically related to an item on the original map or contained a significant part of a correct item label (e.g. "quadrangle" for "croquet lawn" or "union building" for "students union").

The third measure was obtained by asking two Judges to independently rate the maps according to how similar they were with the original map. Judges were told that all the maps came from a single memory experiment, and were asked to concentrate primarily on the spatial information contained in the drawings. Ratings were carried out on a scale of 1 (containing virtually none of the original spatial information) to 10 (containing virtually all of the original spatial information). Inter-rater reliability was high (Spearman's  $Rho = 0.888$ ,  $N = 25$ , two-tailed  $p < 0.0005$ ). The map rating scores were obtained by taking the average rating of the two judges for each drawing.

### *Measures of free recall*

Mann-Whitney U tests were carried out to test for significant differences between the experimental and baseline conditions.

### Free recall of landmark location

Figure 5.9 below shows the mean number of landmarks drawn, the percentage located in the correct or incorrect quadrant and the proportion of errors of commission or unclassified items.



	Number of landmarks drawn	Correct Quadrant	Incorrect Quadrant	Errors of Commission	Unclassified
Experiment	7.88	30.6%	15.0%	9.7%	44.8%
Baseline	12.0	4.2%	18.9%	35.4%	41.7%

**Figure 5.9: Mean scores for the experimental and baseline conditions on a measure of landmark location in free recall.**

Subjects in the baseline condition drew significantly more landmarks than those in the experimental condition ( $U = 13.5$ ,  $N = 25$ ,  $Z = 3.20$ , one-tailed  $p = 0.0007$ ). The experimental group scored more correct quadrant items than the baseline group ( $U = 24$ ,  $N = 25$ ,  $Z = 2.63$ , one-tailed  $p = 0.0043$ ). The experiment group also made significantly fewer errors of commission ( $U = 15$ ,  $N = 25$ ,  $Z = 3.16$ , one-tailed  $p = 0.0008$ ). There were no significant differences between the baseline and experimental conditions on the proportion of incorrect locations or unclassified items. Subjects in the experimental condition drew fewer landmarks, but those landmarks were more likely to be correctly located than would be indicated by use of a guesswork strategy. They also made proportionately fewer errors of commission than would be expected if a guesswork strategy were used.

**Landmark naming**

Figure 5.10 below shows the mean number of landmarks named by subjects, the percentage correctly or incorrectly named and the percentage errors or commission.

	Mean number of landmarks named	Correct Names	Partially Correct Names	Errors of commission
Experiment	7.71	43.5%	33.6%	22.8%
Baseline	12.25	28.6%	16.3%	55.1%

**Figure 5.10: Mean scores for the experimental and baseline condition on a measure of landmark naming.**

Subjects in the baseline condition named significantly more landmarks than those in the experimental condition ( $U = 7.5$ ,  $N = 25$ ,  $Z = 3.55$ , one-tailed  $p = 0.0002$ ). Subjects in the experimental condition produced proportionately more correct names ( $U = 33$ ,  $N = 25$ ,  $Z = 2.04$ , one-tailed  $p = 0.021$ ), partially correct names ( $U = 23.5$ ,  $N = 25$ ,  $Z = 2.61$ , one-tailed  $p = 0.0046$ ) and proportionately fewer errors of commission ( $U = 16.5$ ,  $N = 25$ ,  $Z = 3.03$ , one-tailed  $p = 0.0012$ )



than would be suggested by a guesswork strategy indicated by the performance of the baseline condition.

**Map similarity ratings**

The mean map similarity rating for the experimental condition was 4.8 compared with 2.0 for the baseline condition. This difference was statistically significant ( $U = 12.5$ ,  $N = 25$ ,  $Z = 3.19$ , one-tailed  $p = 0.0007$ ). Independent judges rated maps drawn during free recall by subjects in the experimental condition as being more similar than those of subjects asked to draw a typical university campus.

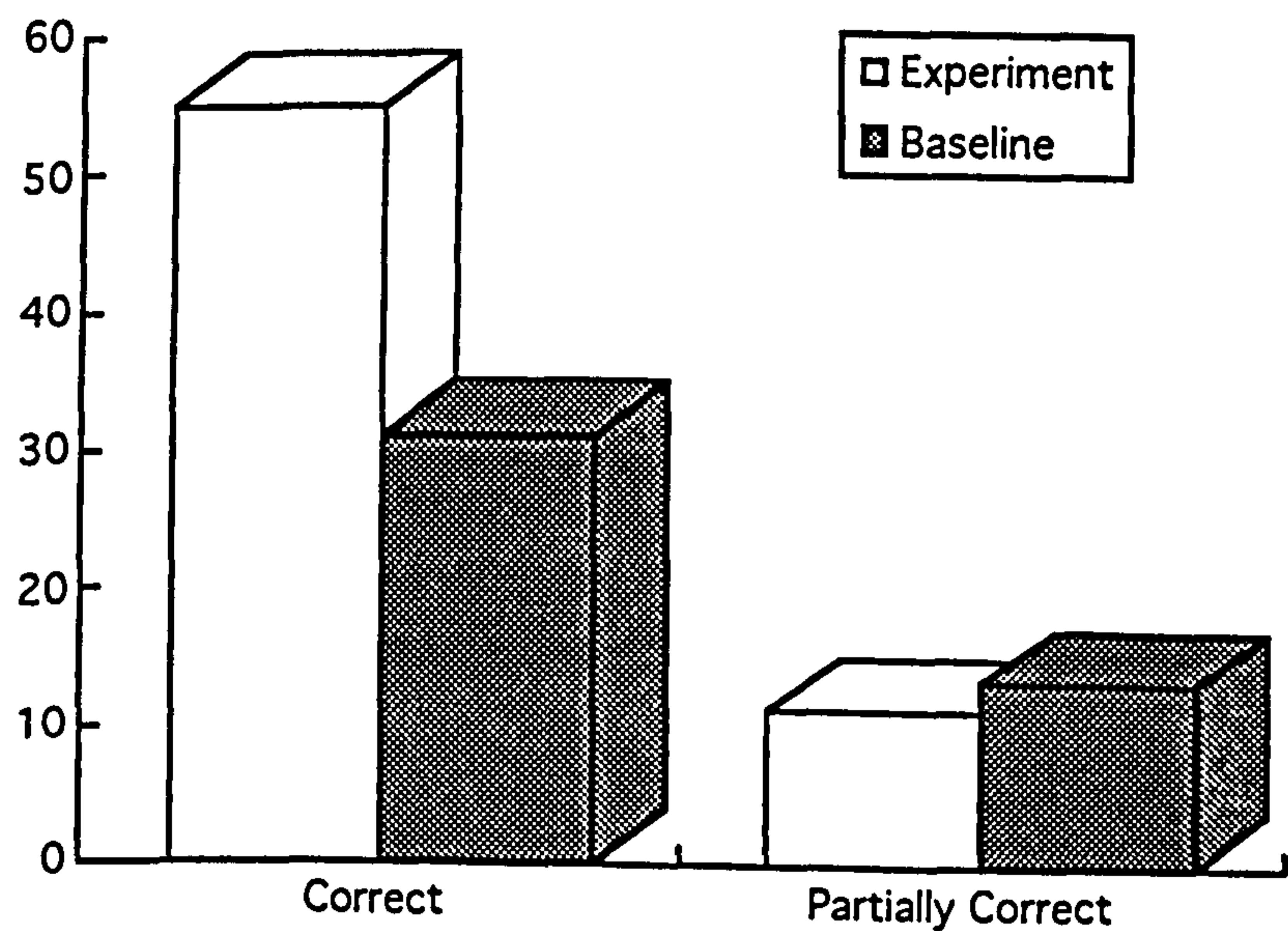
**Place name recognition**

*Scoring procedures*

Landmark names circled on the recognition task were classified as correct, partially correct or incorrect according to the features present on the original map. Partially correct items were those which contained a significant part, but not all, of a feature name on the original map (e.g. 'Union Shop' for 'Students Union') or which were longer or more detailed than the original item (e.g. 'Science Library' for 'Library').

*Recognition performance*

Levels of performance for the experimental and baseline groups on the recognition task are shown in Figure 5.11.



**Figure 5.11: Mean percentage recognition of landmark names.**



Between group differences were analyzed using the Mann-Whitney U test. No significant differences arose for the recognition scores of partially correct items. However, the experimental condition showed significantly higher recognition than the baseline subjects where the landmark names were exactly as presented in the original learning episode ( $U = 12$ ,  $N = 25$ ,  $Z = 3.194$ , one-tailed  $p = 0.0007$ ). This would suggest that partial recall is more readily influenced by background knowledge. Partially correct names may provide extrinsic cues (Jones, 1978; 1982) but recognition under these circumstances may be easily confused with plausibility or familiarity from other contexts (Mandler, 1980). By contrast, exact names provide intrinsic cues for recognition and may reduce confusion by providing unique contextual information.

### Recall of landmark location

#### *Scoring procedures*

Two scoring procedures were used. The first simply recorded the number of correctly placed objects for each subject. The second procedure was a measure (to the nearest half centimetre) of the distance of each object placed on the map from the correct location. Distances were measured from the centre of each of the features located on the map. Using this metric two scores were calculated. The first was the mean error distance taken over thirteen features on the map. The second was the mean error distance for the incorrect items only.

#### *Recall of landmark location*

The mean scores for each subject group on the landmark location task are shown in Figure 5.12. Differences were analyzed using unrelated t-tests.

	Correctly located items	Mean distance error (all items)	Mean distance error (incorrect items)
Experiment	25.5%	7.07 cm	9.12 cm
Baseline	14.4%	7.83 cm	9.22 cm

**Figure 5.12: Mean percentage of correctly located landmarks and mean error in centimetres.**

Subjects who took part in the original experiment scored significantly better on the number of correct object locations than the baseline subjects ( $t = 2.52$ , d.f. = 22, one-tailed  $p = 0.0099$ ). The scores for mean error (taken over all items) showed a similar pattern; experimental subjects



scored lower than baseline subjects ( $t = 2.00$ , d.f. = 22, one-tailed  $p = 0.029$ ). However, there were no significant differences for the mean error distance of the incorrect items.

**Route recall**

*Scoring procedures*

The chief problem in scoring the drawing of the route was that, provided subjects recalled the start and end point, many of the intervening points could be inferred by a simple strategy e.g. taking the most direct route that avoids obstacles. The first measure of route recall was made by scoring the correct start and end locations (these were taken to be either the actual start and end points or the nearest feature to them i.e. the ‘medical centre’ for the start location or the ‘zebra crossing’ into Mandela Park for the end location). The second measure of route recall was based on the number of buildings recalled as being entered during the route. Because not all the intervening buildings on the route were entered it is not possible to guess which buildings were entered by remembering the start and end points of the route.

*Comparisons with baseline subjects*

Means of the recall scores for the route drawing task are presented in the Figure 5.13. Differences were analyzed using Mann-Whitney U tests.

	Start of route	End of route	Mean number of buildings entered	Buildings entered (correct)	Buildings entered (errors of commission)
Experiment	62.5%	62.5%	2.75	86.4%	13.6%
Baseline	0%	0%	2.00	56.3%	43.8%

**Figure 5.13: Mean recall for the start, end and buildings entered on the route.**

Recall of the starting location of the route was significantly higher for the experiment than the baseline group subjects who were at floor ( $U = 24$ ,  $N = 25$ ,  $Z = 2.87$ , one-tailed  $p = 0.0021$ ). Recall of the end point of the route was significantly higher than that achieved by subjects using a guesswork strategy ( $U = 24$ ,  $N = 25$ ,  $Z = 2.87$ , one-tailed  $p = 0.0021$ ). There was no significant difference in the number of buildings entered on the routes drawn by the experimental and baseline subjects. However subjects in the experimental condition identified significantly more correctly entered buildings ( $U = 26$ ,  $N = 25$ ,  $Z = 2.41$ , one-tailed  $p = 0.008$ ) and made significantly fewer errors of commission ( $U = 36.5$ ,  $N = 25$ ,  $Z = 1.90$ , one-tailed  $p = 0.029$ ) than indicated by the guesswork strategy employed by the baseline subjects.



### *Scoring*

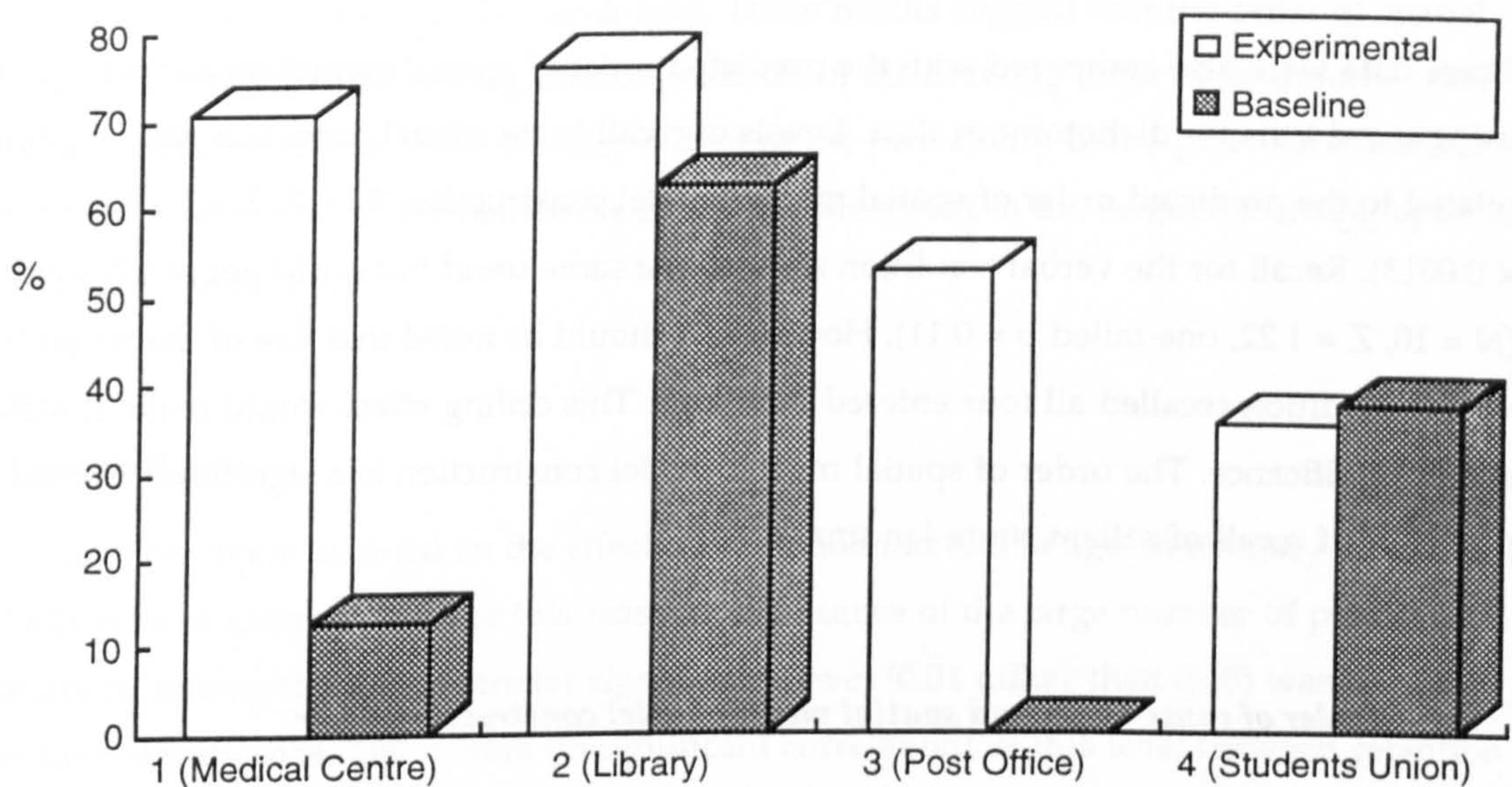
Route recall was compared with the predicted order of spatial mental model construction. The spatial mental model construction order predicts that items integrated early on into a spatial mental model are more likely to be recalled. Unfortunately a simple analysis of recall against the order of landmarks in the route is not possible for two reasons. Firstly, the original route differentiates between buildings which are entered and those which are 'passed by'. It is highly likely that memory for a landmark may differ for those buildings that are entered and those which are not. Secondly, it is not always easy to determine whether or not a building which is not entered lies on any given drawing of the route. For example does a route which passes within one centimetre of the edge of a building or one which passes with two centimetres count as being "on" a route? Restricting the route to buildings actually entered is certainly less arbitrary than determining the route by a fixed distance from the edge or centre of a building. For this reason the analysis is confined to those four buildings which were entered. In order, these are the 'medical centre', the 'library', the 'post office' and the 'students union'.

The spatial model account predicts that the earlier buildings on the route are more likely to be recalled than the later items. It also predicts that the order of route recall should be influenced by the spatial mental model construction. However, the analysis of drawing orders will be influenced by the probability that a given landmark is recalled. This is because when only a single landmark is recalled it will always be drawn first. Previous analyses of drawing order were only carried out when drawing orders were at or close to ceiling and so avoided this problem. An alternative procedure for scoring drawing order which avoids this problem is to classify each landmark as being drawn in the first or last half of the route. If a landmark is drawn exactly in the middle of a route it is excluded from the analysis. This scoring procedure is conservative (it does not use all the available information the data), but is not biased by high or low levels of recall for particular landmarks. For example, if only one landmark is drawn it is excluded from the analysis, because it is neither in the first or last half of the route. If three landmarks are drawn the first landmark is scored as being drawn in the first half of the route and the third landmark as being drawn in the last half of the route (the second landmark is ignored).

### *Probability of route recall and spatial mental model construction order*

Figure 5.14 shows the mean percentage recall of the four entered buildings for the experimental and baseline conditions.

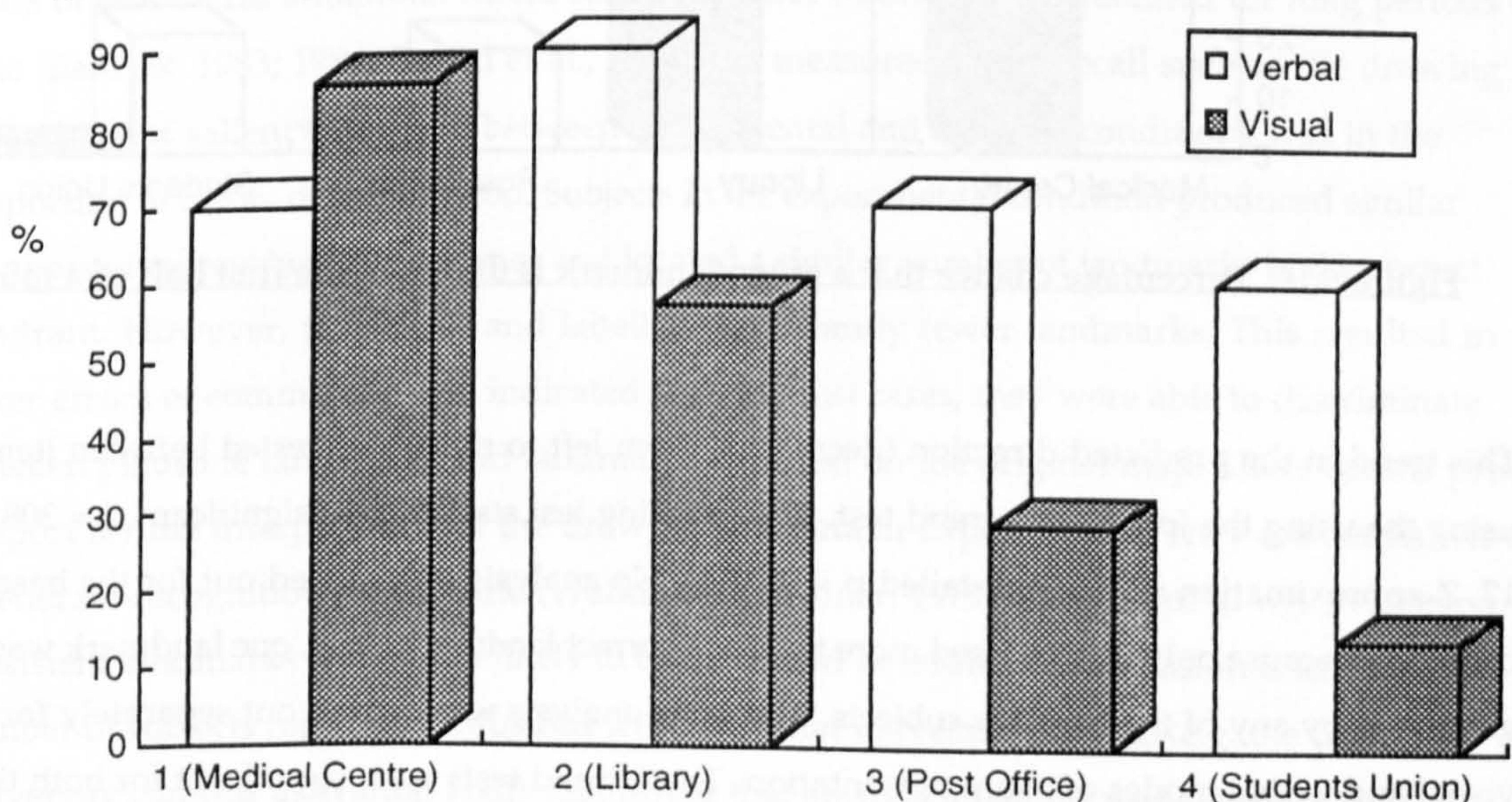




**Figure 5.14: Mean percentage recall for the four buildings entered on the route for the experimental and baseline condition.**

Percentage recall for the four landmarks was compared with the predicted order of spatial mental model construction using a trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). This trend was significant for the experimental condition ( $N = 17$ ,  $Z = 3.01$ , one-tailed  $p < 0.0013$ ) but not for the baseline condition ( $Z = 0$ ,  $N = 8$ , one-tailed  $p = 0.50$ ). This result is consistent with the predicted order of spatial mental model construction. Recall for the library is higher than predicted, though performance in the baseline subjects suggests that this may be inflated by guesswork by some subjects.

Figure 5.15 shows levels of recall for the four buildings entered on the route according to the original mode of presentation of the route (*verbal* or *visual*).



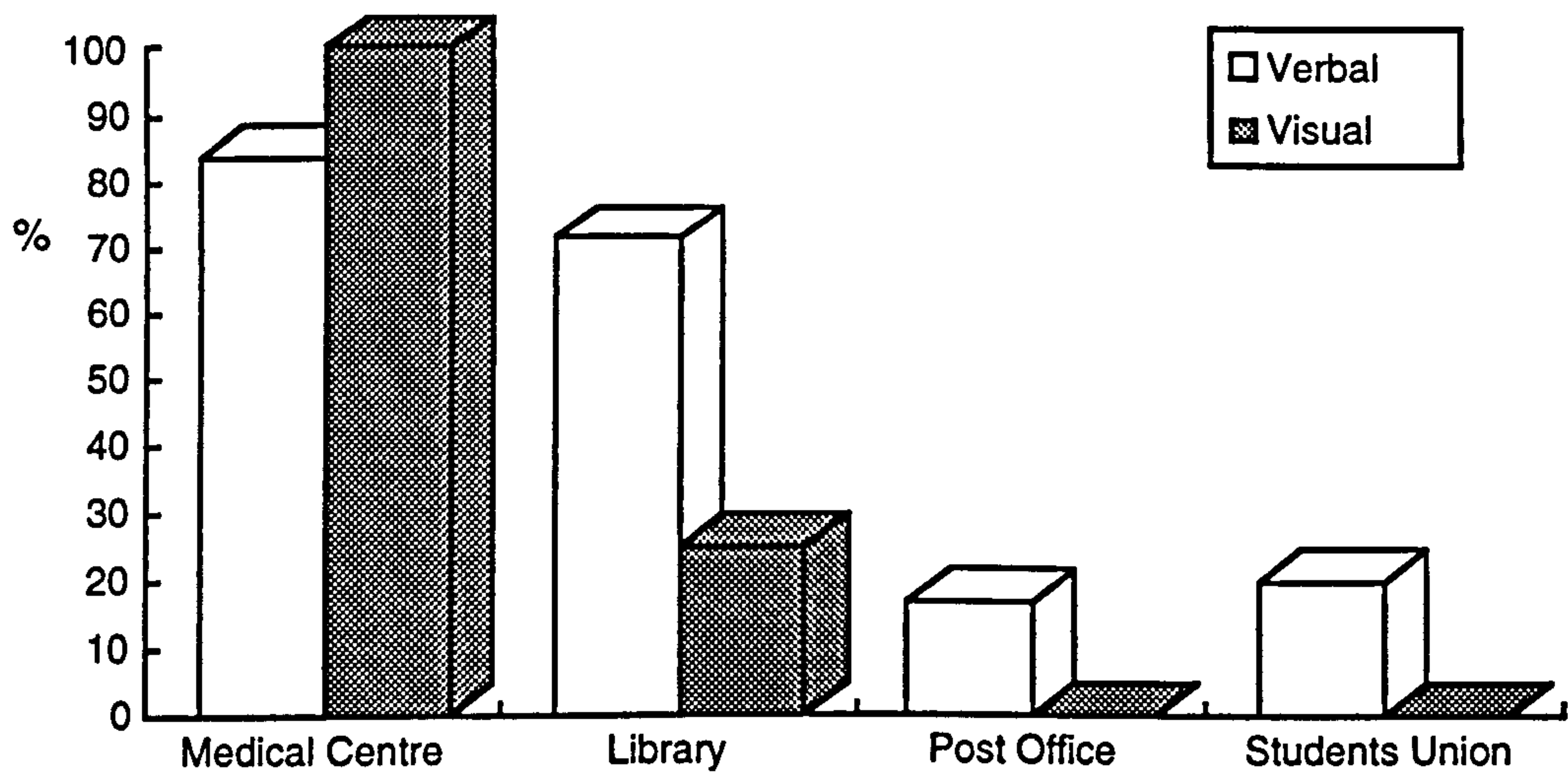
**Figure 5.15: Mean percentage recall for the four buildings entered on the route for the verbal and visual mode of route learning.**



These data were also compared with the predicted order of spatial mental model construction using trend tests for dichotomous data. Levels of recall in the visual condition were significantly related to the predicted order of spatial mental model construction ( $N = 7$ ,  $Z = 3.19$ , one-tailed  $p < 0.0013$ ). Recall for the verbal condition showed the same trend but it did not reach significance ( $N = 10$ ,  $Z = 1.22$ , one-tailed  $p = 0.11$ ). However, it should be noted that five of the subjects in the verbal condition recalled all four entered buildings. This ceiling effect would make it difficult to detect significance. The order of spatial mental model construction is a significant overall predictor of recall of salient route landmarks.

*Order of route recall and spatial mental model construction order*

The position each landmark occupied in the drawing order of the route was classified as falling into the first or last half of the route (see the scoring procedures described above). This scoring procedure ensures that probability of recall and drawing order are not confounded. Otherwise the best remembered landmark, which is more likely to be the only landmark recalled, is more likely to be drawn first. Figure 5.16 shows the percentage chance that an item is drawn in the first, rather than the last, half of the route (scores are depicted separately for the verbal and visual conditions).



**Figure 5.16: Percentage chance that a given landmark is drawn in the first half of a route.**

This trend in the predicted direction (decreasing from left to right) was tested between items using the Jonckheere trend test. The resulting test statistic was significant ( $J = 308$ ,  $N = 17$ ,  $Z$  approximation = 2.95; one-tailed  $p = 0.0016$ ). No analysis was carried out for the baseline condition because only two guessed more than one correct landmark. Also, one landmark was not generated by any of the baseline subjects. The same analysis was carried out separately for verbal and visual modes of route presentation. These trend tests were significant for both the verbal ( $J = 151.5$ ,  $N = 10$ ,  $Z$  approximation = 2.27; one-tailed  $p = 0.012$ ) and the visual ( $J = 27.5$ ,  $N$



= 7,  $Z = 2.12$ ; one-tailed  $p = 0.017$ ) conditions. These results suggest that the order of spatial mental model construction is a significant predictor of the drawing order of landmarks when the route is recalled, and this result is independent of the probability that a given landmark is recalled. Nor does it matter whether the route was presented in the original learning episode in a visual or a verbal form.

### Correlations

No correlations were planned on the effects of retention interval or age on memory for spatial or spatially related information. For this reason and because of the large number of possible correlations among the data a stricter significance level (0.01 rather than 0.05) was used to determine significance. There were no significant correlations at this level between retention interval and measures of recall or recognition. This is not surprising as the effects of retention interval are usually most marked in the first few hours, days or weeks after a learning episode rather than months or years (Baddeley, 1990). Similarly there was no significant correlation of age with most measures of recall or recognition. though there was a trend for older subjects to have poorer free recall for landmark names than younger subjects ( $Rho = -0.653$ ,  $N = 17$ , one-tailed  $p = 0.0057$ ). This is consistent with studies that show that people and place names can be particularly difficult to remember for older adults (Cohen, 1990; Cohen & Faulkner, 1986).

### 5.5.5 Discussion

Well over a year after learning a map of a university campus and a route taken across that map memory is reliably above chance when contrasted with subjects using guesswork strategies. This is particularly striking as the experimental context is relatively impoverished compared to the kinds of real world situations where spatial or other information is retained for long periods of time (Bahrick, 1983; 1984; Cohen et al., 1992). On measures of free recall such as the drawing task the most salient difference between experimental and baseline conditions was in the proportion of errors of commission. Subjects in the experimental condition produced similar numbers of correct landmark names and located a similar number of landmarks in the correct quadrant. However, they drew and labelled significantly fewer landmarks. This resulted in fewer errors of commission and indicated that, in most cases, they were able to discriminate between plausible landmarks and landmarks depicted on the original map. These results provide support for the interpretation of the drawing task data in Experiment 7. They are consistent with a generate-recognition recall route (Watkins & Gardiner, 1979). The extent to which schema consistent landmarks were more likely to be generated as errors of commission is less clear. A number of subjects reported confusions with a familiar university campus (e.g. the Open University campus at Walton Hall) suggesting that episodic memory was also a source of intrusion errors or interference. More recent schema theories have proposed that episodic memory



may be embedded within event or place schema under certain circumstances (Cohen, Kiss, & Le Voi, 1993; Graesser, Gordon, & Sawyer, 1979; Schank, 1982).

The most striking finding is that the order of spatial mental model construction is a significant predictor of both recall of landmarks and, independently, of the order in which those landmarks are drawn. One possible objection to these findings is that the analysis was confined to buildings entered on the route. This was necessary for several reasons. Determining whether a landmark is part of a route in any other way would have entailed adopting arbitrary criteria for inclusion. Using the criteria of entering a building is simple to score and is not arbitrary. Mental models are representations of structural relations (Johnson-Laird, 1983; Johnson-Laird, 1989; Tversky, 1991). Entering a building is a clear example of a spatial and hence structural relationship. Passing or being in the vicinity of a landmark is difficult to quantify in terms of structural relations. Even if it were possible to quantify these relationships there is still a strong case to be made that buildings actually entered on the route are likely to be better remembered than other landmarks through being both the immediate physical and psychological location of the protagonist. Such locations receive greater activation, are better primed and more easily remembered according to several mental models theorists (Bower & Morrow, 1990; de Vega, 1992; Glenberg & Langston, 1992; Glenberg et al., 1987).

The order and probability of recall are both significantly related to the order of spatial mental model construction. Mode of route learning does not appear to be implicated in the order of probability recall or drawing order. Where the original presentation of the route was visual both probability of recall and drawing order are predicted by the order of spatial mental model construction. Where the original presentation of the route was verbal the order landmarks are entered on the route is significantly related to the order of spatial mental model construction. The probability of recall for landmarks in the verbal route learning condition is also consistent with the spatial mental model account though it does not reach significance (however only four subjects contributed to the analysis because of ties). The ordinal relationships obtained for landmark recall on the route drawing task are very difficult to explain in terms of memory for the form the route was originally presented in (memory for what are described as 'comprehension processes' in Chapter 2). There is no *a priori* reason why either a representation of the language of the verbal route should result in superior memory for buildings entered early on in the route than those entered later in the route. Similarly, there is no reason for a visual image of the route to result in better memory for buildings early in the route. Therefore it is unlikely that these results reflect memory for the characteristics of the visually presented route *as well as* memory for the language of the verbally presented route (these issues are discussed in greater detail in Chapter 3). A more parsimonious explanation would be in terms of an amodal representation which preserves the structure of the route (Tversky, 1991). This is consistent with work which shows that people construct an integrated representation of the structure of a situation (i.e. a



mental model) whether it is derived from a text or from a picture story (Gernsbacher, 1985; 1990; 1991).

This experiment has provided the first direct evidence of memory for a spatial mental model at retention intervals of many months rather than hours or days. These findings were obtained with relatively rich and meaningful material albeit with very brief learning periods of minutes rather than weeks or years as in some studies (Bahrick, 1983; Herman et al., 1979). Using such rich materials may provide an opportunity for greater elaboration and improved retention not afforded by simpler materials. Alternatively, simpler materials may not be as greatly influenced by background or schematic knowledge at retention. The next experiment explores the very long term retention of simpler, less meaningful spatial arrangements.



## 5.6 Experiment 9

The two previous experiments in this chapter have reported findings about the very long-term retention of maps or spatial narratives. In this experiment memory for less meaningful spatial information is explored. Subjects from Experiment 4 (reported in Chapter 4) were tested on their recall and recognition for two of the spatial descriptions they had learned at a retention interval of ten to twelve months. The two descriptions chosen were the best remembered determinate and indeterminate descriptions from the original experiment (this was determined by selecting the determinate description with the highest gist recognition and the indeterminate description with the highest verbatim recognition). The two best remembered descriptions were chosen for two reasons. Firstly, it was likely that other descriptions from the original experiment might not be remembered at all. This would make it more difficult to detect small differences in memory. Secondly, descriptions in the original experiment differed markedly in how well they were remembered. This would make it problematic to compare memory for different descriptions if some were reordered and some were not.

Each subject was asked to carry out two tasks. First they were asked to recall any of the names or spatial relations from the two descriptions chosen from the original experiment. These descriptions were identified by their category title ('Landmarks' or 'Condiments'). Secondly, they were given a recognition test similar to the one used in the original experiment. The recognition test contained four different descriptions (the original description, an inferable description and two incorrect descriptions or foils). As before, subjects were asked to rank all four descriptions according to how similar they were to the original description. Half the subjects received the original descriptions in the correct order on the recognition test, while the remainder were given them in a different order. All the descriptions on the recognition test (apart from the original descriptions which had not been reordered) were constructed to have an episodic construction trace overlap of 1 with the original description (Payne, 1993).

The spatial mental model account makes several clear predictions. Firstly, the recall of object names and spatial relationships should reflect the predicted order of spatial mental model construction. In the drawing recall task of Experiment 5 this was confirmed for determinate but not indeterminate descriptions. For indeterminate descriptions recall was better predicted by the frequency an item was mentioned in the original description (see sections 4.2 and 4.2.3). Secondly, memory for the structure of a spatial mental model is indicated if gist recognition exceeds verbatim recognition. The spatial mental model account predicts that gist recognition scores should be higher than verbatim recognition scores for the determinate description. Thirdly, memory for the process of spatial mental model construction (the episodic construction trace) is indicated if reordering reduces verbatim or gist recognition (Payne, 1993). This is because reordering reduces the overlap between the operations used to construct the



original descriptions and the operations used to construct the reordered description. These predictions have been discussed in more detail earlier in the thesis (refer to sections 2.4.5 and 4.1).

The strongest evidence of memory for the structure of a spatial mental model is shown if there is an interaction between the type of recognition and the determinacy of the description. If gist recognition is higher than verbatim for determinate descriptions, but verbatim recognition is higher than gist for indeterminate descriptions this requires a representation of the structure of the situation and a representation of the language of the description (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). Mani and Johnson-Laird have proposed mental models are remembered better because of depth or amount of processing ( Craik & Lockhart, 1972; Craik & Tulving, 1975; Mani & Johnson-Laird, 1982). If this view is correct then the interaction they report may not be present at very long-retention intervals because the level of verbatim recognition for indeterminate descriptions may have fallen to chance.

### 5.6.1 Method

#### Subjects

Twenty-nine subjects who had previously taken part in Experiment 4 (see Chapter 4) took part in this experiment.

#### Materials

Two sets of recognition materials (1 and 2) were created for this experiment. The materials were based on the recognition test materials used in phase 2 of Experiment 4. In set 1 the original descriptions were in the same order that they had been presented in phase 1 of Experiment 4. In set 2 the original descriptions were reordered. Unlike the materials used in Experiment 4 only two descriptions were included in each set. The two descriptions contained in each set were 'Landmarks' (the determinate description with the highest level of gist recognition in Experiment 4) and 'Condiments' (the indeterminate description with the highest level of verbatim recognition). Set 1 thus contained stable versions of the original 'Landmarks' and 'Condiments' description. While set 2 contained reordered versions of the original 'Landmarks' and 'Condiments' description.

As in Experiment 4 three alternative descriptions were presented along with the stable or reordered original description. The three alternative descriptions consisted of an inferable description and two foils. The inferable, foils and reordered original descriptions were constructed to have an episodic construction trace overlap of one with the original description



(Payne, 1993). The inferable and foil sentences shared a single intact sentence with the reordered or stable original descriptions. Both sets of recognition materials were organized so that the original and inferable descriptions were in different positions (labelled A, B, C or D) in the recognition materials.

### **Procedure**

The experiment was divided into two stages. In stage one every subject was asked to try and remember the original learning episode (phase 1 of Experiment 4). They were reminded that the experiment involved reading spatial descriptions off index cards and remembering them for later in the experiment. They were then asked to try and remember two of the descriptions from the original experiment. Half the subjects were asked to remember the 'Condiments' description first. The remaining subjects were asked to recall the 'Landmarks' description first. The only cue provided was the category title of the description. They were told to write down any of the names of the items in the experiment and any of the of the spatial relations between those items that they could remember.

In stage 2 they were reminded that they had been given a recognition test in the original experiment. They were told they would be given a similar recognition test for the two descriptions they had just tried to recall. In each case they were asked to rank four descriptions (labelled A, B, C and D) from 1 to 4 according to how similar they were to the original description they had read. In each case they were informed that one of the descriptions contained all four sentences from the original description, but that those sentences may have been in a different order. They were asked to rank all four descriptions even if they were not certain of the choices they had made. Half the subjects were given the set 1 recognition materials. The remainder were given the set 2 recognition materials. The order of presentation within each set was the same order in which subjects had attempted to recall the two descriptions (i.e. either 'Landmarks' first or 'Condiments' first). This order was counterbalanced with the recognition materials used (set 1 or set 2).

## **5.6.2 Results**

### **Free recall**

#### *Scoring*

Subjects responses were classified according to whether they correctly recalled the names of items from the spatial descriptions they had been presented with in Experiment 4. Where more than one correct name was produced scores were also obtained for the recall of spatial relations



present in the original verbal description. A name was classified as correctly recalled if it contained part or all of the original name or if it was a synonym for that item (e.g. “Mountains” or “Hills” were both acceptable alternatives for ‘Mountain’). Other responses were classified as errors of commission. A spatial relation was classified as correctly recalled if it matched a spatial relation explicitly mentioned in the original description (that is if the names were classified as correct and if the spatial relationship between the two objects was consistent with a sentence in the original description).

<p><b><u>The determinate description (Landmarks):</u></b></p> <p>The city is in front of the mountain.  The city is behind the lake.  The city is to the left of the forest.  The desert is behind the forest.</p> <p><b><u>Inferred spatial relations:</u></b></p> <p>Mountain and Lake.  Mountain and Desert.</p>	<p><b><u>Indeterminate description (Condiments):</u></b></p> <p>The sugar is in front of the ketchup.  The ketchup is to the right of the salt.  The ketchup is to the right of the pepper.  The salt is behind the mustard.</p> <p><b><u>Inferred spatial relation:</u></b></p> <p>Mustard and sugar.</p>
<b><u>Corresponding spatial layout:</u></b>	<b><u>Corresponding spatial layouts:</u></b>
<div> <div>Mountain</div> <div>Desert</div> <div>City</div> <div>Forest</div> <div>Lake</div> </div>	<div> <div>PepperSaltKetchup</div> <div>MustardSugar</div> <div>or</div> <div>SaltPepperKetchup</div> <div>MustardSugar</div> </div>

**Figure 5.17: Spatial relations between objects in the two descriptions recalled in Experiment 9.**

Errors of commission in the recall of spatial relations are not reported because the proportion of errors of commission increases as more correct object names are recalled. For example, if one correct spatial relationship is recalled and four correct names this could result in one correct spatial relation being scored and three errors of commission. This is problematic because someone who recalled fewer correct items would result in fewer errors of commission. Instead, inverse spatial relations are reported (i.e. those occasions where a spatial relation is recalled which is the exact inverse of an actual spatial relation). For example someone who recalled that “The Forest was to the left of the City” would be scored as an inverse spatial relation (because it was actually to the right of the city). In addition to the recall of the relations explicitly mentioned in the spatial descriptions it is also possible to score recall for inferred spatial relations. Figure 5.17 depicts the spatial relationships between objects in the



determinate and indeterminate description used in this experiment (objects in italics are connected by an inferable spatial relation). Inferred relations are true of the description but not explicitly mentioned in it. There are two such inferred relations for a determinate description and only one for an indeterminate description.<sup>4</sup>

*Free recall of names*

Scores were obtained for the percentage recall of all five items from each of the two target descriptions. Errors of commission were also noted. Figure 5.18 shows number of items recalled and the percentage of correct items and the proportion of errors of commission for the determinate ('Landmarks') and indeterminate descriptions ('Condiments'). The median number of errors of commission is also shown.

	Number of item names	Percentage correct	Percentage errors of commission	Errors of commission (median)
Determinate	2.276	65.2%	34.8%	0
Indeterminate	2.793	91.4%	8.6%	0

**Figure 5.18: Number of items recalled and the percentage correct and errors of commission for the determinate and indeterminate descriptions.**

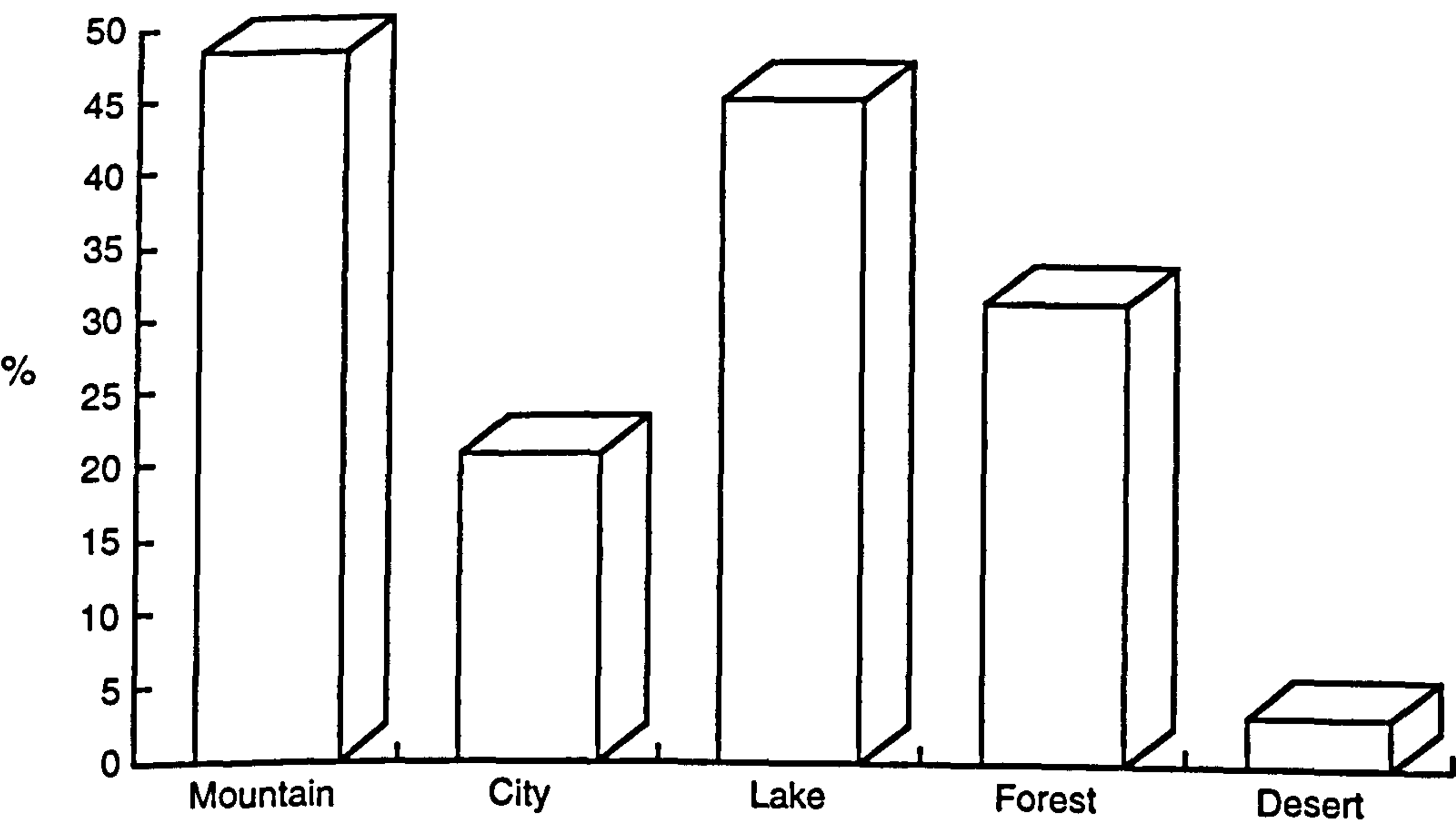
The proportion of errors of commission for the determinate description appears to be very high. However, as the median number of errors of commission shows, this figure is inflated by a few subjects who make large numbers of errors of commission. If the analysis is confined to those subjects who recall at least one correct item from the determinate description the proportion of errors of commission drops to 14.%. Thus it appears that those subjects who are not remembering *any* items from the original experiment are generating most of the errors of commission. This may be because they are not able to recall any items or perhaps because they have been unable to identify the original study episode. This effect is more pronounced for the 'Landmarks' description than the 'Condiments' description. This probably reflects the large number of common landmarks and the small number of common condiments not actually used in the descriptions (in fact only one common condiment "vinegar" was ever produced as an error of commission). Where

<sup>4</sup> This is not strictly true, because it is also possible to infer relationships which are not simple relations such as 'front', 'behind', 'left' or 'right'. However, including diagonal relations such as 'in front and to the left hand side of' would complicate the analysis considerably. It would also be particularly difficult for the indeterminate description where more than one diagonal relation could be inferred between the same two objects.



subjects are able to recall at least one correct item the proportion of errors of commission is very small.

It was predicted that the probability that a given item is recalled should reflect its position in the order of spatial mental model construction. Figure 5.19 shows the probability of recall by the predicted order of spatial mental model construction for the determinate 'Landmarks' description. The trend predicted by the spatial mental models account is decreasing from left to right. These data were analyzed using a non-parametric trend test for dichotomous data (Marascuilo & McSweeney, 1967; Meddis, 1984). The predicted trend was significant ( $N = 29$ ,  $Z = 2.890$ ; one-tailed  $p = 0.0019$ ). This pattern cannot be accounted for in terms of the frequency that items were mentioned in the original description because 'City' was mentioned three times, 'Forest' twice and 'Mountain' only once (the exact opposite of the actual pattern of recall). Nor can frequency of usage explain the observed levels of recall; 'City' is the most frequent item followed, in order, by 'Forest', 'Lake', 'Mountain' and 'Desert' (Kucera & Francis, 1967). Therefore it appears that the order of spatial mental model construction does predict the probability that an item from a determinate description is recalled up to a year after the original learning episode.<sup>5</sup>

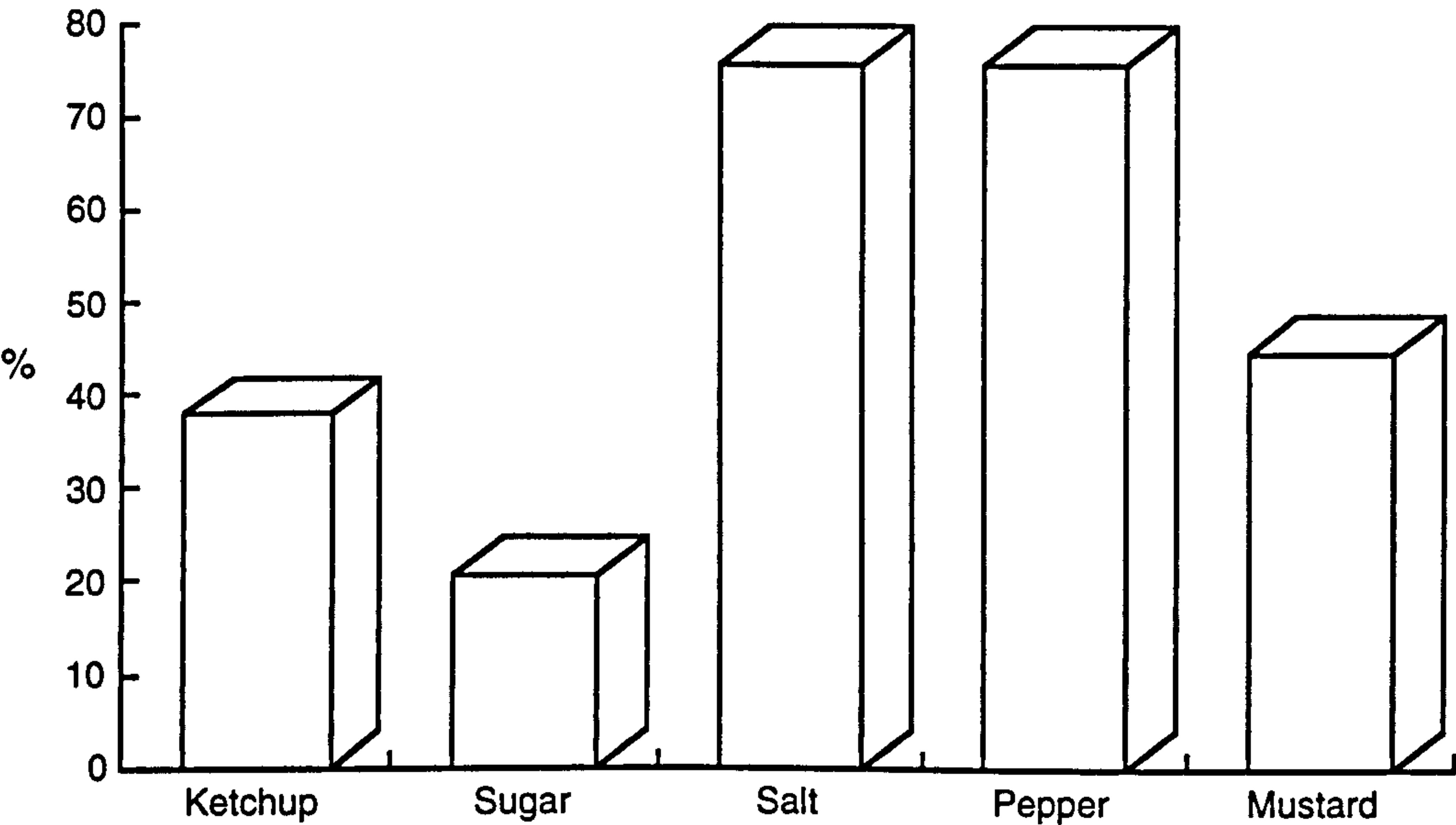


**Figure 5.19: Percentage recall for items in the determinate description (the predicted trend is decreasing from left to right).**

<sup>5</sup> Only memory for 'City' does not conform to the predicted trend, however, a number of apparent errors of commission such as "Tower", "Building" and "Castle" were classified as incorrect even though they might be considered plausible alternatives for 'City'.



The same prediction was also tested for the indeterminate description. Figure 5.20 shows the mean percentage recall for each item in the indeterminate description. The predicted order of spatial mental model construction is from left to right.



**Figure 5.20: Percentage recall for items in the indeterminate description (the predicted trend is decreasing from left to right).**

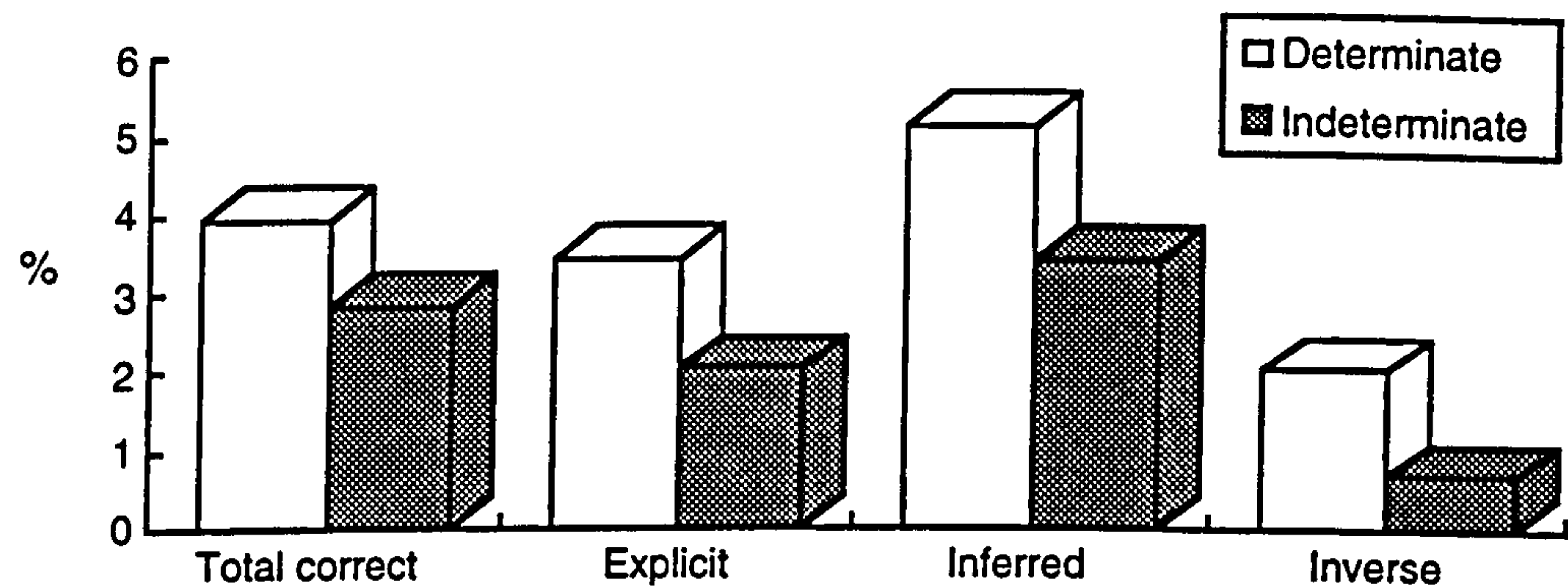
From Figure 5.20 it is clear that, if anything, probability of recall is negatively related to the predicted order of spatial mental model construction. Analysis of this trend shows that this negative relationship is close to significance ( $N = 29$ ,  $Z = 2.697$ ; two-tailed  $p = 0.070$ ). Unlike in the drawing recall task of Experiment 5 the number of times an item is mentioned in the original description is not a good predictor of recall; ‘Ketchup’ occurred three times, ‘Salt’ twice and ‘Pepper’, ‘Sugar’ and ‘Mustard’ only once. Word frequency was also considered as a possible influence on recall (Kucera & Francis, 1967). However, it is not a significant predictor of recall ( $N = 29$ ,  $Z = 0.81$ , one-tailed  $p = 0.21$ ). A plausible alternative is that the pattern of recall is predicted by the degree to which an item is a good exemplar or the target category (this could also be interpreted as the extent to which the items are consistent with a ‘Condiments’ schema). This hypothesis would be consistent with excellent memory for ‘Pepper’ and ‘Salt’ and poor memory for ‘Sugar’. It would also be consistent with the generate-recognize recall route that was proposed to account for free recall data in Experiments 7 and 8 (Jones, 1978; Watkins & Gardiner, 1979). As in Experiment 5 there is no evidence of a positive relationship between the order of spatial mental model construction and probability of recall for indeterminate descriptions.

*Free recall of spatial relations*

Figure 5.21 depicts the mean percentage of correctly recalled spatial relations for determinate and indeterminate descriptions out of the possible total of seven for determinate and six for



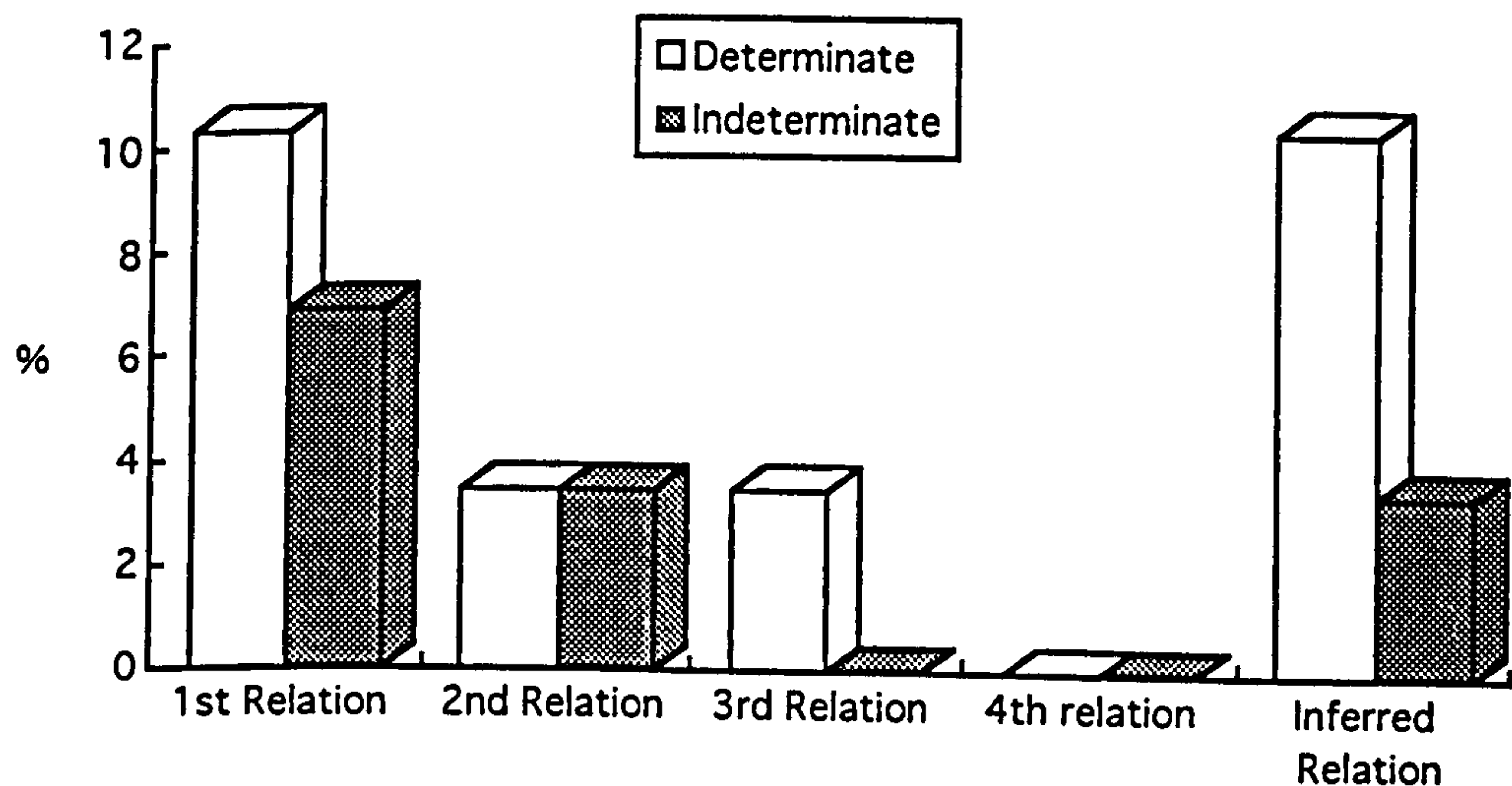
indeterminate descriptions. Also shown are the proportion of explicit, inferred or inverse spatial relations which were generated by subjects.



**Figure 5.21: Mean percentage recall of correct spatial relations (the total is broken down for explicit and inferred relations) and mean percentage of inverse relations recalled.**

Percentage recall for spatial relations is very low. This is because the majority of subjects reported that they could not recall any spatial relations. More correct spatial relations were recalled than inverse spatial relations. However, this did not reach significance with the Fisher exact test (Siegel & Castellan, 1988) either for the total correct ( $N = 29$ , one-tailed  $p = 0.14$ ) or for the inferred spatial relations ( $N = 29$ , one-tailed  $p = 0.087$ ).

Figure 5.22 shows the percentage recall of the four explicit spatial relations for both determinate and indeterminate descriptions. The trend predicted by the spatial mental model account is decreasing from left to right. Also depicted is the best remembered inferred relation.



**Figure 5.22: Mean percentage recall of explicit and inferred spatial relations for determinate and indeterminate descriptions.**



The data for the recall of explicit spatial relations were analyzed using a trend test for dichotomous data. The trend predicted by the spatial mental model account was significant for both the determinate ( $N = 29$ ,  $Z = 2.00$ , one-tailed  $p = 0.024$ ) and the indeterminate descriptions ( $N = 29$ ,  $Z = 2.24$ ; one-tailed  $p = 0.013$ ). These results match those obtained for the recall of spatial relations in Experiment 5. It is possible that the trend for the determinate description may be an artefact of the high recall for early items in the spatial mental model construction order. However, this explanation could not account for the trend observed for the indeterminate description.<sup>6</sup> Therefore there is an indication that the probability of recall for these spatial relations conforms to the order predicted by the spatial mental model account for both determinate and indeterminate descriptions. This suggests that some fragments of information about the spatial relations in a mental model constructed in working memory may be present in long-term memory many months after the original learning episode.

From Figure 5.22 it is clear that the best remembered inferred spatial relation is more likely to be recalled than the fourth spatial relation taken from the original description. This would be particularly difficult to account for in terms of memory for an episodic construction trace or memory for the language of the text. However, with the Fisher exact test this difference does not reach significance ( $N = 29$ , one-tailed  $p = 0.059$ ).<sup>7</sup>

## Recognition

### *Scoring*

As in Experiments 4 and 6b recognition was scored using the procedure described by Payne (1993). The chi square analyses carried out in this section require that gist and verbatim recognition scores be independent (see Experiment 4 and Experiment 6b). Where this was necessary, cases

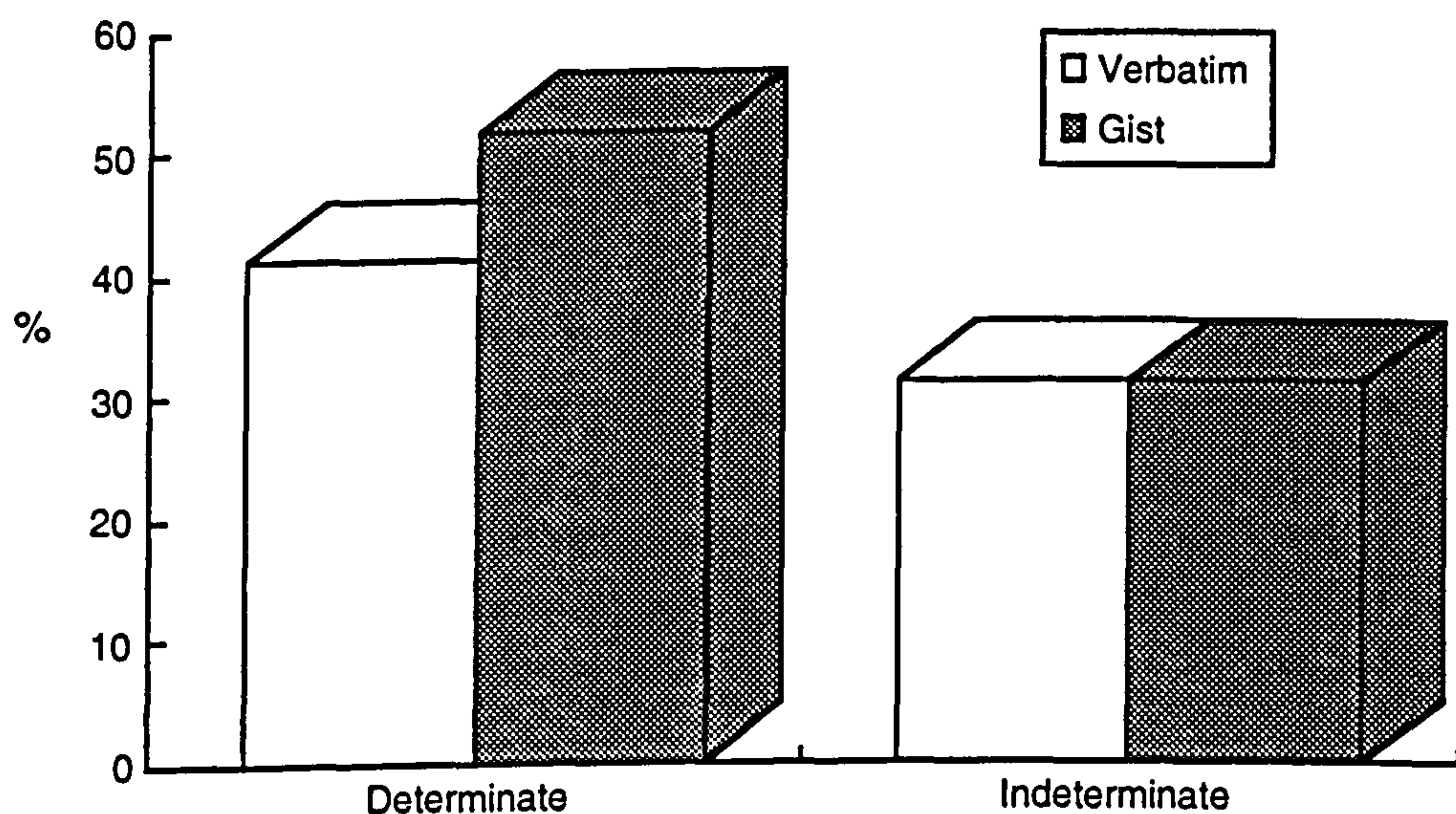
---

<sup>6</sup> Note that the only explicit spatial relations remembered from the indeterminate description are from the first and second spatial relations. The indeterminacy of this particular description only arises after the third spatial relation has been read. The inferred spatial relation is of course true with respect to either of the possible situations described by the indeterminate description. This result is consistent with the spatial mental model account.

<sup>7</sup> It should be noted that because very few spatial relations were recalled it was possible that memory for inferred spatial relations was independent of memory for explicit spatial relations. When the raw scores were checked this was discovered to be the case. In no case was an inferred spatial relation recalled only because two explicit spatial relations which formed the premises of the appropriate spatial inference were recalled. Where inferred and explicit spatial relations are not independent (as in Experiment 5) it would be inappropriate to report recall for inferred spatial relations. This is because if all four explicit spatial relations are recalled all the possible inferences will also be 'recalled'.



where a given subject was scored for both gist and verbatim recognition were not included in the analysis. For the remaining analyses and for the summary statistics these cases are included. However, as these cases contribute only to 'tied' scores they would not influence the pairwise comparisons carried out in the following sections.



**Figure 5.23: Mean percentage verbatim and gist recognition for the two descriptions.**

### *Is recognition above chance?*

Figure 5.23 shows mean percentage gist and verbatim recognition scores for the determinate and indeterminate description. Chance recognition performance is 25% for verbatim recognition and 16.7% for gist recognition. These levels of recognition were tested against chance performance using the binomial test (Siegel & Castellan, 1988). For the determinate description verbatim recognition is above chance but this level does not reach significance ( $N = 29$ ,  $p = 0.082$ ). The probability of determinate gist recognition exceeding fifty percent, as it does, is significant ( $N = 29$ ,  $p = 0.000016$ ).

However, it could be argued that gist recognition may be inflated by verbatim recognition performance. One possible check on this is to analyze only those cases where gist and verbatim recognition occur independently of each other. In these cases verbatim recognition is at 19.0% ( $N = 29$ ,  $p = 0.83$ ). Gist recognition in the absence of verbatim recognition is at 33.3% ( $N = 29$ ,  $p = 0.082$ ). This estimate of significance is highly conservative because it assumes that *every* case where gist and verbatim recognition occur together arises only because the subject notices that the inferable description describes the same situation as his or her first choice. In other words it assumes that where gist and verbatim recognition occur together true gist memory is *never* present. Determinate verbatim recognition is lower than determinate gist recognition and is not significantly above chance (on either a conservative or a lenient test). A more plausible



interpretation is that gist recognition is not exaggerated by verbatim recognition. In fact the data suggest the reverse; above chance gist recognition has inflated determinate verbatim recognition.

Indeterminate gist and verbatim recognition scores were analyzed in the same way. At 31% indeterminate verbatim recognition is not significantly above chance ( $N = 29$ ,  $p = 0.28$ ). Gist recognition of the same level is above chance ( $N = 29$ ,  $p = 0.041$ ). Using the more conservative estimate of significance described above neither verbatim ( $N = 29$ ,  $p = 0.89$ ) nor gist recognition ( $N = 29$ ,  $p = 0.58$ ) is significantly above chance. These data suggest that indeterminate verbatim recognition is not significantly greater than chance. Overall, indeterminate gist recognition is greater than chance, though it is possible that verbatim recognition scores may have inflated gist recognition.

Overall recognition scores are higher than can be accounted for by chance. There is no evidence that verbatim recognition is above chance for either determinate or indeterminate descriptions. By contrast there is evidence that gist recognition is above chance for both determinate and indeterminate descriptions. These results support the view that the structure of a mental model is better remembered than the sentences it is derived from, perhaps because of deeper encoding or more elaboration ( Craik & Tulving, 1975; Johnson-Laird, 1983; Mani & Johnson-Laird, 1982; Tardieu et al., 1992).

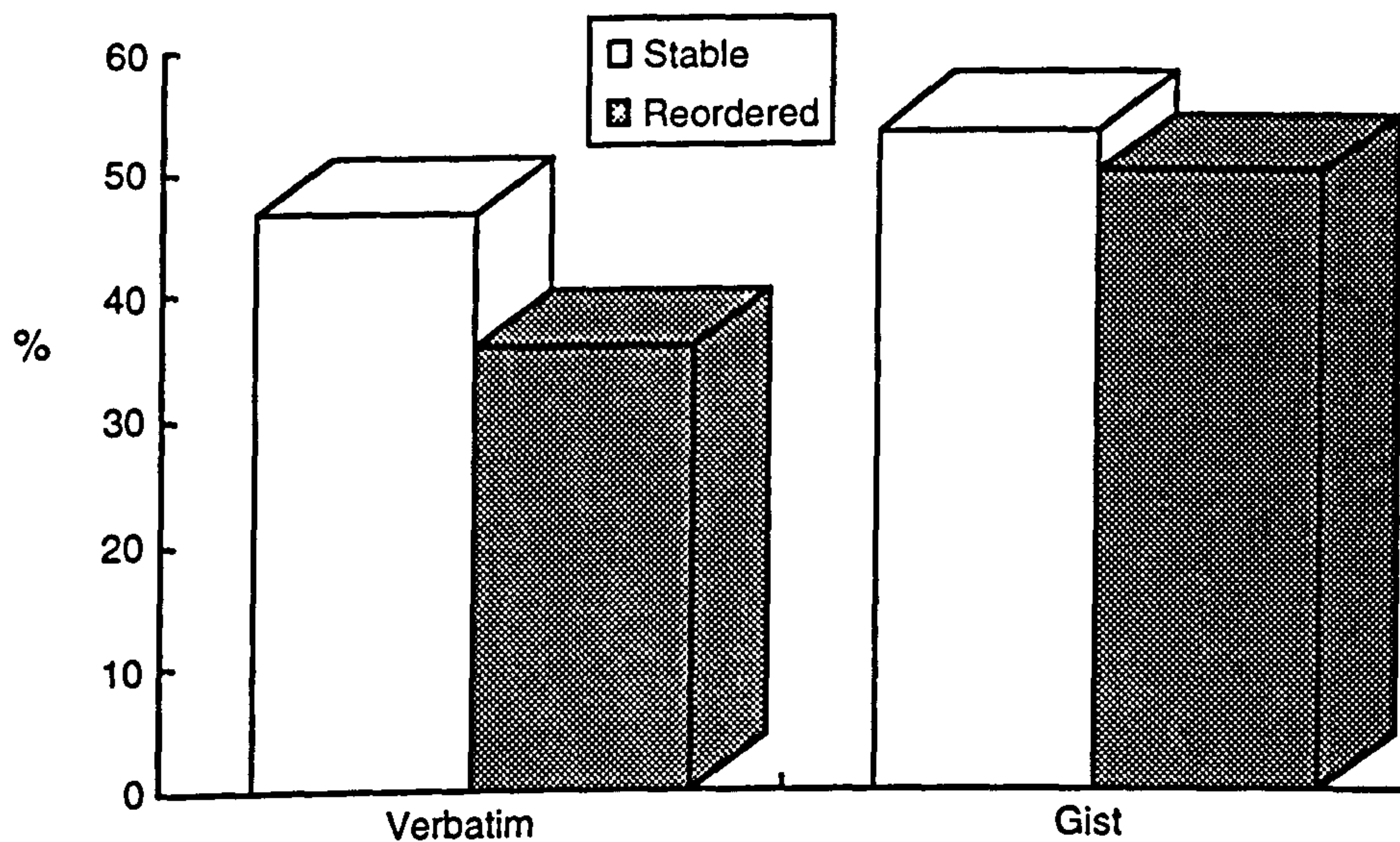
#### *Is there an interaction between determinacy and gist or verbatim recognition?*

There is little indication of an interaction between gist and verbatim recognition and the determinacy of the description in Figure 5.23. It is probable that any interaction may have been eliminated through the effect of the long retention interval on verbatim recognition performance. The data in Figure 5.23 were analyzed using the Fisher exact test (the chi square test was not used because of low expected cell frequencies). As described in the scoring section above only independent observations were included in this analysis. The resulting statistic was not significant ( $N = 29$ , one-tailed  $p = 0.48$ ). At very long retention intervals there is no evidence of an interaction between determinacy and type of recall.

#### *Does reordering depress recognition memory?*

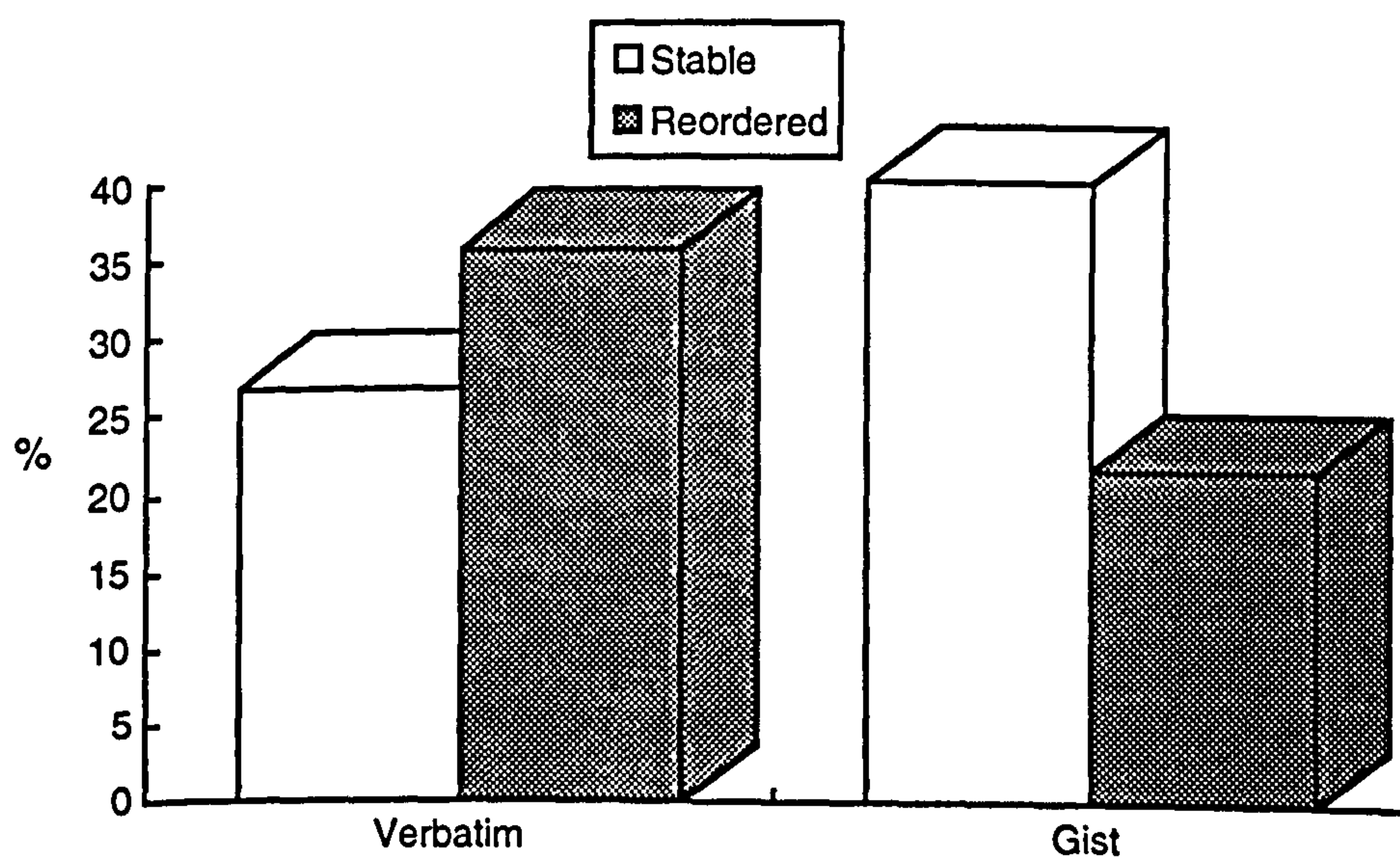
Figure 5.24 depicts the effect of reordering on gist and verbatim recognition for the determinate description only. Reordering depresses both verbatim and gist recognition slightly. This reduction is not significant for either verbatim recognition ( $U = 93.5$ ,  $N = 29$ ,  $Z = 0.59$ , one-tailed  $p = 0.28$ ), or gist recognition ( $U = 101.5$ ,  $N = 29$ ,  $Z = 0.18$ , one-tailed  $p = 0.43$ ) when analyzed with the Mann-Whitney U test.





**Figure 5.24: The effect of reordering on mean percentage gist and verbatim recognition (determinate description only).**

Figure 5.25 depicts the effect of reordering on gist and verbatim recognition for the indeterminate description only. Reordering does not depress verbatim recognition for the indeterminate description at all (the difference between the two scores is not significant). However, this may be because verbatim recognition is at or about chance. Gist recognition is depressed by verbatim recognition, though this effect is not significant ( $U = 85.5$ ,  $N = 29$ ,  $Z = 1.06$ , one-tailed  $p = 0.14$ ).



**Figure 5.25: The effect of reordering on mean percentage gist and verbatim recognition (indeterminate description only).**

Reordering does not significantly depress either gist or verbatim recognition scores. These results do not favour the hypothesis that recognition is based on memory for an episodic construction trace (Payne, 1993). It is possible that some subjects retain memory for the



construction of a spatial mental model. Larger sample sizes may find support for this view. However, these results would suggest that at long retention intervals memory for an episodic construction trace is poorer than that for the structure of a spatial mental model.

### *Correlations with backward digit span*

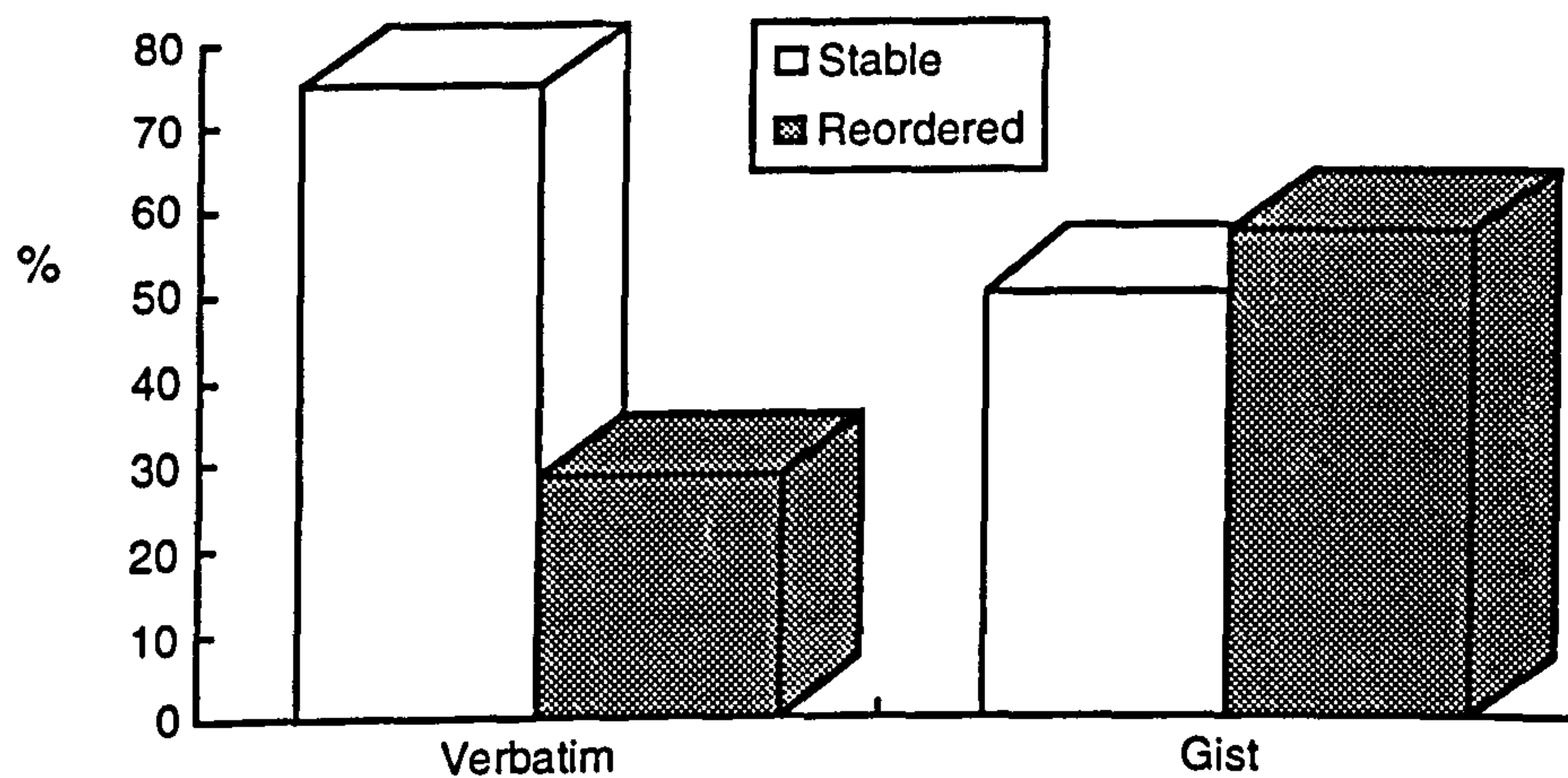
In the original study an association was observed between backward digit span and gist recognition scores, but no correlation was shown between reading span and either gist or verbatim recognition. The data in this experiment were also correlated with backward digit span using the Spearman rank correlation coefficient. Backward digit span was not significantly correlated with either gist ( $Rho = 0.14$ ,  $N = 29$ ,  $Z = 0.075$ , two-tailed  $p = 0.45$ ) or verbatim recognition ( $Rho = 0.007$ ,  $N = 29$ ,  $Z = 0.036$ , two-tailed  $p = 0.97$ ).

### *Correlations with retention interval*

Retention interval was correlated with subjects' ranks for the original and inferable descriptions in the recognition task. For the determinate description retention intervals ranging from 308 to 368 days were significantly correlated with ranks assigned to the original description ( $Rho = 0.448$ ,  $N = 29$ ,  $Z = 2.36$ , one-tailed  $p = 0.0092$ ) but not for ranks assigned to the inferable description ( $Rho = -0.01$ ,  $N = 29$ ,  $Z = -0.055$ , one-tailed  $p = 0.48$ ). This suggests that verbatim recognition for the determinate description is significantly poorer at longer retention intervals. Gist recognition, however, remains stable over these retention intervals (roughly ten to twelve months). For the indeterminate description neither the ranking of the original ( $Rho = 0.128$ ,  $N = 29$ ,  $Z = 0.68$ , one-tailed  $p = 0.25$ ) nor of the inferable ( $Rho = -0.229$ ,  $N = 29$ ,  $Z = -1.21$ , one-tailed  $p = 0.112$ ) is significantly associated with retention interval.

Closer examination of the data revealed that retention interval and determinate verbatim recognition are correlated only for the stable description ( $Rho = 0.666$ ,  $N = 15$ ,  $Z = 2.49$ , one-tailed  $p = 0.0064$ ). Figure 5.26 shows determinate verbatim and gist recognition scores for stable and reordered descriptions at median or below median retention intervals. Reordering more than halves verbatim recognition in these cases. This effect was significant when analyzed with the Mann-Whitney test ( $U = 43$ ,  $N = 15$ ,  $Z = 1.74$ , one-tailed  $p = 0.041$ ). In addition, only for the stable determinate descriptions and only at the shorter retention intervals is verbatim recognition significantly above chance ( $N = 8$ ,  $p = 0.0042$  using the binomial test). Gist recognition is not disrupted by reordering. This suggests that some subjects may retain memory for the construction of a spatial mental model in the form of an episodic construction trace. However, this is limited to shorter retention intervals and only for determinate descriptions.





**Figure 5.26: The effect of reordering on mean percentage gist and verbatim recognition at short retention intervals (determinate description only).**

### 5.6.3 Discussion

Recall of objects and spatial relations from the determinate description confirms the predictions of the spatial mental models account. The probability that a given object or spatial relation is recalled is predicted by the order in which items would be integrated into a spatial mental model. For indeterminate descriptions, only the recall of spatial relations is predicted by the order of spatial mental model construction. This may be because it is possible to construct a single, consistent mental model for the first two spatial relations in the indeterminate description. This finding replicates the results obtained in Experiment 5 at longer retention intervals.

Neither the number of times a given item occurs in the original description, nor the frequency of the word in common usage (Kucera & Francis, 1967) is a significant predictor of recall for either determinate or indeterminate descriptions. These results suggest that recall is based on memory for the structure of a spatial mental model. Both an episodic construction trace or memory for the sentences of the text would predict that an item mentioned in more than one proposition is more likely to be recalled. It is not possible to conclude that either an episodic construction trace (Payne, 1993) or a textbase (Kintsch, 1988) are implicated in the recall of determinate or indeterminate spatial descriptions. Recall for the determinate description can be accounted for by memory for the structure of a spatial mental model. Recall for the indeterminate description may be best accounted for in terms of a generate-recognition recall route (Jones, 1978; Watkins & Gardiner, 1979). The items in the indeterminate description were drawn from a category ("Condiments") with relatively few common members. It is possible that high recall for the indeterminate description largely reflects how easily category members are generated rather than 'true' recall.



There was little evidence of the interaction between type of recognition and the determinacy of the description predicted by Mani and Johnson-Laird (1982). The interaction present in Experiment 4 may have been eliminated by poor verbatim recognition performance for both descriptions. Overall verbatim recognition scores were not significantly different from chance for either the determinate or indeterminate description. There was evidence, though, that gist recognition was above chance for the determinate description and possibly also for the indeterminate description. These findings support the hypothesis that the structure of a spatial mental model is better remembered at long retention intervals than either the process of its construction or the language of the original description. This suggests that mental models may indeed be encoded more deeply or elaborated more during the initial learning episode (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). There is certainly evidence that mental models take longer to construct than representations of the language of the text (Tardieu et al., 1992).

Unlike Experiment 4 there is no significant relationship between backward digit span and gist recognition. This suggests that verbal short term memory may aid in the elaboration or construction of a spatial mental model but that any such effect is relatively small or relatively short-lived. Even though overall determinate verbatim recognition is not significantly above chance there is a significant relationship between retention interval and the rank assigned to the original description. As there is no relationship between the rank assigned to the inferable description and retention interval this suggests that at shorter retention intervals (slightly over ten months) some subjects retain some memory for the language of the determinate description. Closer analysis revealed that this relationship holds only when the original description has not been reordered. These results provide some support for the view that verbatim recognition scores for determinate descriptions were influenced by memory for an episodic construction trace. The correlations between stable verbatim recognition and retention interval are difficult to explain purely in terms of memory for the language of the text or for the structure of a spatial mental model. However, memory for the process of mental model construction appears to be less durable than memory for the structure of a spatial mental model.

At very long retention intervals there is no evidence that people remember the language of the spatial descriptions they have read. There is some indication that memory for the process of spatial mental model construction influences determinate verbatim recognition. Recall for items from the descriptions and the spatial relations between them as well as high levels of gist recognition all support the view that people are able to remember the structure of a spatial mental model up to a year after the initial learning episode.



## 5.7 General Discussion

Three experiments are reported which investigate the role of memory for spatial mental models in the very long-term retention of information from spatial descriptions. In the following sections issues raised in the introduction to this chapter and during the discussion of the individual experiments are discussed.

### 5.7.1 Assessing level and type of recall

In the introduction methodological issues relating to the recall of information from long-term memory were discussed. A particular problem for studies of very long-term retention is the role of schema in remembering and especially reconstructing the original learning episode. In Experiment 7 highly typical (schema consistent) buildings were much more likely to be remembered when subjects attempted to recall an English village on a drawing task. Because errors of commission were infrequent (even high typicality errors such as school or garage) it was suggested that recall was consistent with subjects taking advantage of a generate-recognition recall route. High typicality landmarks are more likely to be generated. Once a typical landmark is generated the low proportion of errors of commission suggests that subjects appear to be able to recognize many of the landmarks which were actually present in the original story (Bahrick, 1970; Jones, 1982; Watkins & Gardiner, 1979).

In Experiment 8 recall and recognition performance of subjects in the experimental condition was shown to exceed that of baseline subjects using guesswork strategies. On most of the recall tasks subjects in the experimental and baseline conditions 'recalled' similar numbers of correct items, but baseline subjects produced very many more 'errors'. Again, this is consistent with the generate-recognition recall route. If experimental subjects were able to make greater use of direct or *intrinsic* recall routes this should be reflected in higher raw scores, not just in the proportion of correct responses. Intrinsic recall routes use retrieval cues encoded during the original study episode while indirect or extrinsic routes (such as generate-recognition) use cues that were not encoded during learning (Jones, 1978; 1982; 1987). These results do not exclude the use of direct or intrinsic recall routes by subjects, but they do suggest that indirect recall routes dominate subject responses at very long retention intervals.

The interpretation of recall performance in Experiment 9 is less clear cut. For the determinate description the trends in the recall of items and spatial relations suggests that many subjects were able to use the description title "Landmarks" as a direct retrieval cue for information from a spatial mental model. The correct retrieval of items may also have provided direct access to some of the remaining landmarks. According to the spatial mental model account



items in a description are integrated into a consistent representation in working memory (Johnson-Laird, 1983). Hence the spatial mental model account should predict that landmarks in the description act as intrinsic cues at recall and that the effectiveness of these cues should reflect the order in which items were integrated into the representation (Jones, 1978; 1982). Recall of the indeterminate descriptions suggests that indirect, extrinsic cues may be more important in recalling individual items (though this may not be the case with the spatial relations). This is consistent with the view that subjects are unable or at least less likely to integrate indeterminate descriptions into a single consistent representation at encoding. However, this interpretation is weakened by the fact that the 'Condiments' category for the indeterminate description has a smaller set size than the 'Landmarks' category. The high levels of 'recall' for the 'Condiments' description probably reflects the difficulty of generating condiments which were not in the original description.

### 5.7.2 What kinds of information are preserved?

The findings reported here are broadly consistent with previous research on memory at very long retention intervals. Research on the retention of spatial information (Bahrick, 1983; Herman et al., 1987; Herman et al., 1979) suggests that what is remembered reflects the needs and goals people have when learning an environment. In a similar fashion, subjects in these studies who were originally asked to concentrate on the spatial information in the descriptions they read, appeared to retain spatial information better than the language of the descriptions. It has been argued that the retention of this spatial information reflects, to a great extent, memory for the structure of a spatial mental model.

Bahrick also noted that the more quickly information about an environment was learned the more quickly it was later forgotten (Bahrick, 1983). This is consistent with the evidence that subjects retain some memory for an episodic construction trace in Experiment 9. The episodic construction trace is presumably encoded at an early stage in the construction of the spatial mental model. Further elaboration of a spatial mental model may aid the retention of its structure but not its construction. The episodic construction trace, therefore, is more sensitive to the effects of retention interval. This is shown by the correlation between stable verbatim recognition and retention interval for the determinate description. Where the original description was reordered retention interval does not influence recognition. Gist recognition, which reflects memory for the structure of a spatial mental model, is stable over these retention intervals. This is likely to be because of deeper encoding or elaboration of the structure of a spatial mental model during encoding ( Craik & Tulving, 1975; Mani & Johnson-Laird, 1982).



More generally the results have supported the view that schema can have a dramatic influence on recall (Cohen, 1989). In these experiments the influence of schematic knowledge appears to be largely during retrieval. Schema consistent items tend to be more likely to be generated during recall, but the low proportion of errors of commission suggests that, in most cases, subjects are able to reject plausible but incorrect items. This is consistent with experiments that show that a change in schema at retrieval can influence recall (Anderson & Pichert, 1978; Pichert & Anderson, 1977). Schema theories cannot explain all of the information that subjects recall. They cannot explain memory for routes, memory for spatial location, spatial relations or the relationship between probability of recall and the predicted order of spatial mental model construction. What these things have in common is that they can be integrated into a single coherent mental representation during the original learning episode. In fact, some of these, such as spatial relations, are properties of an integrated mental representation.

## 5.8 Conclusions

Evidence has been presented from three studies of the very long-term retention of spatial information. The findings presented in this chapter primarily reflect two kinds of processes; the influence of schemas on retrieval from long-term memory and the retention of the structure of a spatial mental model. It is proposed that the notion of an integrated mental representation such as a mental model is a useful one for exploring the retention of spatial information at very long retention intervals.



## Chapter 6 - Conclusions

In the first two chapters of this thesis it was proposed that to answer the question 'Are spatial mental models represented in long term memory?' requires three important assumptions. Firstly, that it is necessary or useful to postulate at least three forms of 'high-level' representation in working memory (Anderson, 1983; Johnson-Laird, 1983). Secondly, that one form of high-level, dynamic representation in working memory, namely mental models, preserves the "relation-structure" of the situations it represents (Bower & Morrow, 1990; Craik, 1943; Johnson-Laird, 1983; Tversky, 1991; van Dijk & Kintsch, 1983). These 'mental models' can be distinguished from representations of language, in the form of linguistic propositions, or representations of the perceptual characteristics of situations, such as visual images (Johnson-Laird, 1983). Thirdly, that the processes in working memory can profoundly influence how something is represented in long-term memory (Craik, 1979; Craik & Lockhart, 1972; Marschark & Surian, 1989; Morris et al., 1977).

Spatial mental models are of special interest to cognitive scientists because the characteristics of the situations they describe are relatively easy to manipulate and because they "happen to be a domain in which the differences between the structure of the world and the structure of the text describing it are most readily apparent" (Garnham & Oakhill, 1994). Accordingly, the nine experiments presented in this thesis have focused on memory for the structure of spatial descriptions. In earlier chapters these experiments have, to some extent, stood alone. In this chapter the findings from these experiments will be considered together and discussed in terms of the spatial mental modeling processes outlined in Chapter 2.



## 6.1 Spatial mental models and memory for spatial descriptions

According to the framework outlined in Chapter 2 mental modeling processes can be divided into three stages (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Payne, 1993). These stages are considered in terms of the special case where a spatial mental model is derived from a verbal description. In the first stage, *comprehension*, the meaning of the spatial relationships in the description is accessed and interpreted. In the second stage, *construction*, tokens corresponding to real or imagined entities are placed into the spatial mental model according to the possible interpretations of the spatial relationships in the description. Where possible all new tokens are added relative to existing tokens in the spatial mental model. In the final stage, *consultation*, information is read off from the spatial mental model or more complex operations such as matching or counting are carried out using the representation (Glenberg & Langston, 1992; Johnson-Laird, 1983; Newell, 1990). Comprehension processes may operate differently or be omitted where a spatial mental model is derived from perception or previous experience. Construction processes may also operate differently where a spatial mental model is recalled from long-term memory. In addition construction processes are also involved in updating spatial mental models during the comprehension of complex discourse or when searching for counterexamples in deductive reasoning (Bower & Morrow, 1990; Byrne & Johnson-Laird, 1989; Glenberg et al., 1987). Consultation processes are involved in monitoring the construction of the spatial mental model as well as in generating inferences during discourse comprehension and reasoning (Byrne & Johnson-Laird, 1989; Johnson-Laird, 1993; Johnson-Laird & Byrne, 1991).

These three stages of processing operate in a cyclic fashion. In the simplest case the acquisition of each new spatial relation requires comprehension, construction and consultation to operate in series. In more complex cases mental modeling may require that comprehension and consultation processes are repeated in order to check the interpretation of the spatial mental model is consistent with the situation described. If the interpretation is considered inappropriate, construction processes may have to be altered (e.g. by deleting a token from the spatial mental model). In extreme cases construction may even have to be abandoned (Mani & Johnson-Laird, 1982; Payne, 1993). It is the cyclic nature of mental modeling that supports many of the predictions of memory for spatial descriptions made in this thesis (see section 2.5). Empirical support for these predictions has already been discussed in detail in Chapters 3, 4 and 5. The main findings which have emerged from these experiments are summarized in the sections which follow.



### 6.1.1 Are drawing orders influenced by the order of spatial mental model construction?

Several experiments reported in this thesis have addressed the question of whether the order in which landmarks are drawn in a drawing task is a result of memory for the order of spatial mental model construction. The main recall task was to draw the environment described in the story, concentrating on the physical and spatial features contained in it. In similar experiments by Taylor and Tversky (1992a) people tended to draw landmarks from the description in the order those landmarks had been introduced in the text. Taylor and Tversky interpreted this finding in terms of subjects reproducing a spatial mental model of the text in the order that it had originally been constructed. However, they failed to address the possibility that people recall a representation of the language of the text. In order to differentiate memory for the language of a text and memory for a spatial mental model several of the experiments reported have investigated memory for scrambled texts. When a text is scrambled the predicted order of spatial mental model construction and the order landmarks are mentioned in the text differ (this argument is discussed in more detail in Chapter 3). The main findings from these experiments are listed below:

- When subjects are asked to recall spatial information from an unscrambled story drawing orders are significantly related to the predicted order of spatial mental model construction (Experiments 1, 2, 3 and 8).
- When sentences from a story describing a route are presented in a scrambled order subjects' drawings are also significantly related to the order of spatial mental model construction (Experiments 2 and 3).
- The first landmark drawn by subjects in a drawing recall task is predicted by the order of spatial mental model construction. This is the case even when the first landmark entered into a spatial mental model is not the first landmark mentioned in the original story (Experiments 2 and 3).
- When subjects read determinate spatial descriptions (descriptions which have no narrative content) drawing orders are also consistent with the order of spatial mental model construction (Experiments 5).
- A significant relationship between drawing orders and the order of spatial mental model construction can still be obtained at retention intervals of over ten months (Experiments 8 and 9).

These findings offer strong support for the view that the order in which landmarks or other objects are recalled in drawing tasks is determined by the order of spatial mental model construction. Where the predicted order of spatial mental model construction and the order



landmarks are mentioned in a story differ the evidence consistently favours the spatial mental model account. There is no evidence, in these experiments, that drawing recall is based on memory for the language of a text.

### **6.1.2 Is the probability of recall in a drawing task influenced by the order of spatial mental model construction?**

Craik (1970) has shown that negative recency effects arise in memory for lists in delayed, but not immediate, recall. One explanation for this phenomenon is that primacy effects in free recall are the product of greater 'semantic' processing for early items in a list (Craik, 1970; Craik & Lockhart, 1972). Craik proposed that people actively try to transfer early items to long-term memory in order to free up working memory capacity for later items. The spatial mental model account outlined in Chapter 2 makes a similar prediction, but for slightly different reasons. It is argued that when items are integrated into a spatial mental model, early items are elaborated to a greater extent than later items. This is because later items are integrated into the spatial mental model by relating them to items already present in the representation. Evidence for a relationship between order of mention and the probability of recall for items from a spatial text has been reported by Denhière and Denis (1988). When it is possible to form a single, coherent representation of the structure of a story, the item mentioned first in a written narrative or picture story takes longer to process and is quicker to access (Gernsbacher, 1990; 1991). Several experiments reported in this thesis have investigated the probability of recalling items from a text or description to test this prediction. Of particular interest is the probability of recall from a text when the order items are mentioned in the text and the predicted order of spatial mental model construction differ:

- When subjects are asked to recall spatial information from an unscrambled route text the probability of recall is significantly related to the predicted order of spatial mental model construction (Experiment 8).
- When subjects read scrambled route texts the probability of recall is also significantly related to the predicted order of spatial mental model construction (Experiments 2 and 3).
- When subjects read determinate spatial descriptions the probability of recalling an object or a spatial relation from a description is significantly related to the order of spatial mental model construction (Experiments 5 and 9).
- When subjects read indeterminate spatial descriptions the probability of recalling a spatial relation (but not the probability of recalling an object) is predicted by the order of spatial mental model construction (Experiments 5 and 9).



- The probability of recalling an object from an indeterminate spatial description is predicted by the number of times that object is mentioned in the original description (Experiment 5).
- A significant relationship between the probability of recall and the order of spatial mental model construction can still be obtained at retention intervals of over ten months (Experiments 8 and 9).

These results support the prediction that the probability of recalling an item from a spatial description or text is influenced by the order in which that item was integrated into a spatial mental model. Where text order and the order of spatial mental model construction differ the results favour the spatial mental model account. One interesting finding is that the relationship between spatial mental model construction order and the probability of recalling an object is not obtained for indeterminate descriptions. This result is problematic for an explanation of order effects in recall purely in terms of memory for construction processes (e.g. an episodic construction trace). This issue is explored further in section 6.1.5 below (also refer to section 4.2.3).

### 6.1.3 Does reordering a spatial description impair recognition memory?

Payne (1993) has argued that memory for the construction of a spatial mental model takes the form of an *episodic construction trace*. An episodic construction trace is a set of propositions where each proposition encodes a single operation of an attempt to construct a spatial mental model. Payne has argued that if a spatial description is reordered for a recognition test this reduces the overlap between the original description and the reordered description, making recognition based on an episodic construction trace more difficult. Reordering the descriptions leaves each sentence intact and so should not impair memory for the sentences of the description. In addition, the reordered description still describes the same arrangement of objects, so reordering should not make recognizing the structure of a spatial mental model more difficult. A number of experiments in this thesis have addressed the question of whether reordering a spatial description does impair recognition memory:

- When subjects hear a series of spatial descriptions reordering the descriptions in a recognition test significantly reduces recognition memory for both indeterminate and determinate descriptions (Experiment 6b).
- When subjects read a series of spatial descriptions some, but not all, descriptions, show a significant effect of reordering (Experiments 4 and 6b).
- There is evidence that reordering significantly impairs recognition memory even after a ten month retention interval (Experiment 9).



These findings are broadly consistent with the predictions made by the episodic construction trace hypothesis . The different effects of reordering when a description is read and when it is heard are interesting. They may arise, at least in part, because people who read a written description are able to deviate from the actual order of presentation (e.g. when the description is re-read).

#### 6.1.4 Do people remember the structure of a spatial mental model?

Memory for the structure of a spatial mental model is demonstrated when people remember both the explicit and implicit spatial relations present in a spatial mental model. One way to demonstrate this is to show that, when conditions favour the construction of a single, determinate spatial mental model, the structure or 'gist' of the model is remembered better than the description itself. Even stronger evidence in favour of the spatial mental model account is obtained if it is demonstrated that people remember the language of the description when conditions do not favour the construction of a single, determinate spatial mental model (e.g. when the spatial description is radically indeterminate). The classic mental models account thus predicts an interaction between determinacy of a description and the type of recognition it supports (Johnson-Laird, 1983; Mani & Johnson-Laird, 1982). Several experiments in this thesis have addressed this issue. The main findings are reported here:

- Gist recognition scores for determinate descriptions are significantly higher than those obtained for indeterminate descriptions (Experiment 4).
- Verbatim recognition scores for indeterminate descriptions are significantly higher than those obtained for determinate descriptions (Experiment 4).
- When the indeterminate and determinate descriptions are controlled for differences between items this interaction between determinacy and type of recognition supported by a description is still obtained for descriptions which have not been reordered (Experiment 5).
- Reordering indeterminate descriptions on a recognition test can reduce verbatim recognition almost to chance (Experiment 6b).

These results suggest that people do remember the structure of a spatial mental model constructed from a determinate description, because determinate gist recognition scores are consistently higher than those for indeterminate descriptions. However, the evidence does not favour the conclusion that the language of the description is always remembered better for indeterminate descriptions than determinate descriptions. This is because reordering an indeterminate



description can reduce verbatim recognition to a level barely above chance. Instead, it suggests that people remember their attempt to construct a spatial mental model for an indeterminate description.

### 6.1.5 Discussion

In Chapter 2 three processing stages involved in spatial mental modeling were outlined. Of the three stages, *comprehension* is the least important for the purposes of this thesis. This is because any theory of how spatial descriptions are understood and remembered must propose some form of linguistic processing stage as its starting point (Clark, 1969; Johnson-Laird, 1981; van Dijk & Kintsch, 1983). Consequently, memory for comprehension processes provides no means of discriminating between mental model and alternative accounts of the comprehension of spatial descriptions. For this reason the main focus of this thesis has been on memory for mental modeling processes of construction and consultation. However, it has been possible to reject explanations which propose people only remember the language of the descriptions or texts they have read. Explanations based purely on memory for the language of a description have problems accounting for several of the findings reported in this thesis. Firstly, it is difficult to explain the relationship between the order or the probability of recalling an item in a drawing task and the order of spatial mental model construction. This is particularly difficult where spatial mental model construction order and the order items are mentioned in the text differ, as is the case with scrambled texts (van Dijk & Kintsch, 1983). Secondly, verbatim recognition scores are lower than gist recognition for determinate descriptions. This is impossible to account for in terms of memory for the language of a description, because it means that people are more likely to 'recognize' descriptions containing only one sentence from the original description than those containing four sentences from the original description. The final finding against people remembering only the language of a description is that reordering makes it harder to recognize a description, even though all of the individual sentences remain intact.

Memory for the processes of spatial mental model *construction* and *consultation* is central to the account of mental modeling processes outlined in Chapter 2. Previous mental models approaches have not required that people remember constructing a mental model. In general, they rely on evidence from processing in working memory. For example, evidence that reasoning problems which require multiple models are more difficult than those which require a single model (Johnson-Laird & Byrne, 1989; Johnson-Laird & Byrne, 1991; Johnson-Laird et al., 1992). Demonstrating memory for the processes of mental model construction is important because it confirms detailed predictions about how that theory is implemented. One recent paper has proposed that the process of mental model construction may be the only thing which is retained in long-term memory (Payne, 1993). Payne's proposal that people retain an *episodic construction*



*trace* of the operations used to construct a spatial mental model provides the basis for the account of memory for construction processes proposed earlier in the thesis (see section 2.5.1).

When people are asked to draw spatial layouts learned from spatial descriptions the order in which items are drawn and the probability that an item is recalled are both influenced by the predicted order of spatial mental model construction. If simple spatial descriptions are reordered at test, recognition memory is impaired, suggesting that people remember the operations used to construct a spatial mental model. Explanations which do not rely on memory for spatial mental modeling processes have difficulty accounting for these observations (Payne, 1993; Tversky, 1991). On face value both these observations seem to reflect the retention of information related to the processes of spatial mental model construction such as an episodic construction trace. However, a more detailed examination of this interpretation suggests that the current formulation of the episodic construction trace hypothesis (Payne, 1993) has problems accounting for some of the results reported here. Payne (personal communication) has noted that minor modifications to the episodic construction trace account are required to take account of the relationship between drawing recall and the order of spatial mental model construction. These might include determining the order in which the first two items are entered into the model and assuming that the propositions are not stored in an 'all or none' fashion. A more fundamental problem for the episodic construction trace hypothesis comes from differences in drawing recall between determinate and indeterminate descriptions in Experiment 5. For determinate descriptions the probability of recalling an item was significantly related to the order of spatial mental model construction. For indeterminate descriptions this was not the case. The episodic construction trace records all the operations in the attempt to construct spatial mental models from both determinate and indeterminate descriptions (Payne, 1993). It is therefore difficult to explain differences between memory for determinate and indeterminate descriptions by the episodic construction trace hypothesis.

One possible resolution of the problems facing the episodic construction trace hypothesis is that the idea that people remember the process of spatial mental model construction should be abandoned. This resolution would be unsatisfactory for two reasons. Firstly, reordering a spatial description can make it significantly harder to recognize. This finding is very difficult to account for, except in terms of memory for the operations used to construct a spatial mental model (Payne, 1993). Secondly, there is considerable evidence that people can and do remember the structure of a spatial mental model. If people remember the structure of a spatial mental model, it is not unreasonable to expect them to be able to remember some or all of the operations involved in its construction. For these reasons, the resolution favoured here is to explain memory for the order of spatial mental model construction in terms of the combination of construction and consultation processes (see section 2.5). A 'processing' account of memory for spatial mental models along these lines is proposed in section 6.2.



## 6.2 A processing account of the representation of spatial mental models in long-term memory

In Chapter 2 it was proposed that mental modeling consists of three processing stages; comprehension, construction and consultation. In this section one kind of theory of memory for spatial mental models is rejected and a sketch of a processing model of memory for spatial descriptions is described.

### 6.2.1 Copy theories of memory for spatial mental models

A 'copy' theory of memory for spatial mental models assumes that once a mental model is constructed in working memory it is somehow copied as a complete, intact entity into long-term memory. While no explicit 'copy' theory of memory for mental models has been proposed it could be argued that it is implicit in a number of treatments of memory for mental models. One reading of the paper by Mani and Johnson-Laird (1982) is that mental models share the same format in long-term memory as in working memory, implying that they are copies of completed working memory mental models (Payne, 1993). Other experiments which do not explicitly distinguish between the working memory and long-term memory representations of subjects can also be interpreted in this fashion (Franklin et al., 1992; Tversky, 1991). A simple copy theory is unattractive for several reasons. Firstly, as has already been noted, there is a great deal of evidence to suggest that there are multiple forms of representation in working memory, but relatively little evidence for this in long-term memory (Anderson, 1983; Denis, 1991; Glenberg & Langston, 1992; Kosslyn, 1980; Payne, 1993). Secondly, it seems strange to represent only the final, completed mental model in long-term memory. Mental modeling is often an ongoing activity and even with a static mental model it is not always obvious when the model is complete. Finally, what is stored in long-term memory can be distorted or altered over time by organized knowledge structures in long-term memory (Anderson & Pichert, 1978; Cohen et al., 1993; Schank, 1982). Experiments presented in this thesis have shown the influence schemas can have on recall. This suggests that spatial representations in long-term memory are in sufficiently similar format for them to be influenced by or confused with other long-term memory traces. For this reason any 'copy' theory of memory for mental models has to take into account the nature of the working memory processes used to construct and consult the representation. A plausible version of such a copy theory might consist of 'snapshots' of the mental model construction and consultation processes rather than merely a copy of the completed representation. Such a revised copy theory would be difficult to distinguish from the processing account proposed here.



It could be argued that a simple copy theory such as the one rejected above is unlikely to be advanced as a theory of memory for spatial descriptions. However a similar theory of memory for images, namely dual-coding theory, has been proposed (Paivio, 1971; 1986; 1991). One recent formulation of dual-coding theory could be considered as an alternative explanation of memory for spatial mental models. Kulhavy has described a version of dual coding theory which acknowledges that images have both visual (feature) and spatial (structural) components (Kulhavy, Stock, Peterson, Pridemore, & Klein, 1992; Kulhavy et al., 1993; Kulhavy et al., 1993; Kulhavy et al., 1993). It is the structural components of images which lead to better retention and which lead to pictures or maps improving memory for text. Kulhavy argues that these images are stored as “a more or less unitary chunk” in long-term memory. Furthermore, Kulhavy has suggested that these images (derived from pictures or maps) are not built up sequentially (Kulhavy et al., 1993). This account shares many predictions with the mental models account because both stress the role of structural information in improving retention. Where the two accounts differ is on how that structure arises. The Kulhavy account cannot explain why the predicted order of spatial mental model construction influences recall, both here and in other research (Denhière & Denis, 1988; Taylor & Tversky, 1992a). It is also difficult to reconcile with evidence that ‘images’ derived from perception are built up sequentially rather than copied intact into working memory or long-term memory (Kosslyn, 1980; Kosslyn, Cave, Provost, & von Gierke, 1988).<sup>1</sup> While there is some evidence that such a dual coding approach may successfully account for structural information in ‘images’ derived from pictures it fails to explain how such images could be derived from descriptions. A simple ‘copy’ theory which does not take account of cycles of processing in long-term memory fails as a theory of how people remember spatial descriptions.

### 6.2.2 A sketch of a processing theory of memory for spatial mental models

A productive theory of memory for spatial mental models needs to combine the attractive features of the episodic construction trace hypothesis with a detailed explanation of how consultation processes influence long-term memory. In this section a sketch of such a theory is provided followed by a discussion of how the sketch needs to be fleshed out. In the theory it is assumed that two long-term memory traces are formed; a *construction* trace and a *consultation* trace. The role of language comprehension processes is deliberately ignored, not because they are unimportant, but because several excellent and detailed theories of text processing and memory for text already exist (Bower & Morrow, 1990; Garnham, 1987; Gerrig, 1988; Kintsch, 1988; van Dijk & Kintsch, 1983). The two traces are composed of arrays of tokens in a specific spatial relationship. They are depicted as arrays for several reasons. The first reason is to stress the

---

<sup>1</sup> In the terms of this thesis these spatial ‘images’ correspond more closely to a spatial mental model than to a visual image.



spatial nature of the working memory processes being recorded in long-term memory. The second is that to express the same information explicitly in, for example, a propositional format would be extremely inefficient. To express the spatial relationship between two tokens would require at least two propositions (e.g. "X is to the right of Y" and "Y is to the left of X"). As the number of tokens in the model increases the number of propositions increases rapidly (e.g. for four tokens at least twelve propositions are required). Thirdly, the consultation trace and the construction trace support the recognition and recall of spatial relations from a spatial mental model. It therefore makes sense to describe them as high-level structured entities rather than as relatively low level propositions which can be used to derive such a higher-order entity.

The general principles of the sketch are simple. Both the construction trace and the consultation trace are recorded in long-term memory as a side effect of conscious, effortful processing in working memory (the construction processes and consultation processes described in section 2.5.1). The construction trace is a record of the operations applied in order to construct a spatial mental model and is largely procedural in nature. The consultation trace is a record of additional processing carried out while the spatial mental model is consulted and is largely declarative in nature. Each trace is considered to have constituent structure such as the tokens present in the model and the spatial or procedural relationships among them. This is an important point, because it is assumed that a token such as "mountain" or "pepper" in the model can be associated with other elements in working memory or long-term memory. These elements can in turn be used to cue or prime retrieval from a construction trace or consultation trace recorded in long-term memory (Bower & Morrow, 1990; Glenberg et al., 1987; McNamara, Halpin, & Hardy, 1992; McNamara, Halpin, & Hardy, 1992; Radvansky et al., 1993; Radvansky & Zacks, 1991; Wilson et al., 1993). While it is assumed that tokens and relations in the trace can differ in activation or trace strength this aspect of the model is not explicitly addressed in the sketch presented here. It should be noted, though, that the activation or trace strengths of tokens and relations is implicit in the arrays of the construction and consultation trace which are stored over time.

### **The construction trace**

The construction trace is closely based on the episodic construction trace proposed by Payne (1993). The first change from the original episodic construction trace account is intended to explicitly distinguish between the first and second tokens entered into a spatial mental model. This simply requires a separate array to be recorded for each token added to the spatial mental model. The second change is more complex. The episodic construction trace records only the minimum amount of information about construction operations required to reconstruct a spatial mental model (Payne, personal communication). It follows that a construction trace might record more information than that preserved in the episodic construction trace. For instance, it could record spatial relationships other than those made explicit in the original description. The



position adopted here is that the construction trace has a probability of recording existing tokens in the trace according to their spatial proximity to the token being added to the model. Any tokens explicitly referred to in the sentence of the original description being read are automatically recorded in the model. This provides a mechanism for the retention of spatial mental model construction order in the construction trace (a mechanism which is not present in the episodic construction trace). A detailed description of spatial mental model construction order is provided in section 2.5.1.

To understand the way such a construction trace might work imagine being given a set of three objects in a bag; a knife, a fork and a spoon. You are asked to place the fork in front of the spoon. You take out the fork and spoon and then place the spoon on the table. Next you place the fork in front of the spoon. You are then asked to place the knife behind the spoon and you do so. Afterwards you may remember each operation quite clearly, but let us consider the final operation. You will almost certainly remember placing the knife behind the spoon. However, it is also reasonable to expect you to remember that by placing the knife behind the spoon you also placed it behind the fork. The construction trace is procedural in the sense that people are remembering how a token was added to a spatial mental model rather than what the resulting spatial relationships were (Rumelhart & Norman, 1983; Ryle, 1949). It is this possibility of noticing other spatial relationships that this version of the construction trace captures. Rather than attempt a detailed explanation of how the construction trace is recorded examples of memory for two simple spatial descriptions are described later in this section.

### **The consultation trace**

The consultation trace records declarative information that is read from a spatial mental model. A spatial mental model is usually consulted for a particular purpose such as making an inference (Byrne & Johnson-Laird, 1989), understanding a story (Bower & Morrow, 1990) or carrying out a set of instructions (Glenberg & Langston, 1992). In the consultation of a spatial mental model no distinction is made between implicit and explicit information. For this reason, consulting a spatial mental model will often result in people 'noticing' spatial relations which are not explicit in the original description or in the operations used to construct it (Glenberg & Langston, 1992). For certain tasks the way people consult a spatial mental model may result in almost exclusively recording information which is explicit in the original description. This is unlikely, however, because constructing a spatial mental model requires effort and people are not likely to do so if they can perform the task by referring to the original description.<sup>2</sup> If anything people

---

<sup>2</sup> Reasoning tasks may be an exception to this general principle. According to Johnson-Laird and Byrne (1991) people do make errors in deduction when the conclusion is explicit in the premises. This may well be because they expect the task to be more difficult than it is and construct one or more mental models of the premises in order to reach a conclusion.



may be more likely to search for and record novel spatial relations from the spatial mental model because it is more useful than information explicit in the original description.

The most important and difficult aspect of describing the consultation trace is to determine what spatial relations are recorded and when. While it is possible to argue that consultation processes follow a fixed order like construction processes this would not be consistent with the many different purposes a spatial mental model can be constructed for. For the present a skeletal outline of the consultation processes required to account for memory for indeterminate descriptions will be described. This will be done by distinguishing between the first and subsequent cycles of the consultation processes. The first cycle is assumed to consist of reading the entire description from start to finish, constructing and consulting the spatial mental model as it is read. The subsequent cycles are assumed to take place after construction has been completed and hence only involve consultation processes. For the sake of simplicity the first cycle of consultation processes are assumed to be devoted entirely to checking the spatial relations present in the spatial mental model are consistent with the original description. If a token is not added to the spatial mental model, for example because the description is indeterminate, it is, of course, not possible to consult the novel spatial relationships that might have become part of the model. In subsequent cycles, for each new token which has been added to the spatial mental model its spatial relationship with immediately adjacent tokens is consulted and recorded in long-term memory. Of particular importance is the first cycle. This is because it is not necessarily the case that reading the description from start to finish only involves reading each sentence once. In fact it is probable that some sentences, particularly early sentences, are read several times in order to check the spatial mental model. For the first cycle of consultation processes it is proposed that spatial relations from early sentences in the description are consulted more often than those later in the description. This is depicted in the two examples which follow by spatial relations which occur early in the construction process being followed by a number in brackets indicating that they may be recorded several times. This multiple recording of some sections of the consultation trace could be considered as multiple representations of these propositions in memory or greater activation or trace strength of some long-term memory contents over others.

### Memory for two simple spatial descriptions

Examples of memory for a simple determinate and a simple indeterminate description are presented here to demonstrate how the construction trace and consultation trace are recorded. As noted earlier each trace is made up of a number of 'arrays' consisting of tokens and spatial relations between tokens. In the construction trace the following notation is also adopted. Question marks precede any token for which it is not possible to provide a fixed spatial location in the array. Bold lettering indicates a token which has just been added to the spatial mental model. Normal lettering indicates a token which is explicitly provided as a reference point in the original description (and which is automatically recorded in the trace). Tokens in italics are



those which have a probability of being recorded in the construction trace, but which are not explicitly provided as reference points in the original description. The two descriptions used are shown in Figure 6.1.

<u>Determinate description:</u>	<u>Indeterminate description:</u>
The knife is in front of the fork	The duck is in front of the frog
The plate is to the right of the fork	The mouse is to the right of the frog
The spoon is to the left of the fork	The cat is to the right of the frog
The cup is in front of the spoon	The dog is in front of the cat

Figure 6.1: The determinate and indeterminate descriptions used in Figures 6.2 and 6.3.

Figure 6.2 below depicts a construction trace and consultation trace for the determinate description. Figure 6.3 depicts a construction trace and consultation trace for the indeterminate description. In both cases only the construction trace and the first cycle of the consultation trace is depicted.

Construction operations	Construction trace	Consultation trace (First Cycle)
1. Enter fork in model	Fork	Fork
2. Place knife in front of fork	Fork Knife	Fork (4) Knife (4)
3. Place plate to right of fork	Fork      Plate Knife	Fork      Plate ( 3 )      ( 3 )
4. Place spoon to left of fork	Spoon   Fork   Plate Knife	Spoon      Fork ( 2 )      ( 2 )
5. Place cup in front of spoon	Spoon   Fork   Plate Cup      Knife	Spoon (1) Cup (1)

Figure 6.2: An example of a construction trace and the first cycle of a consultation trace for the determinate description in Figure 6.1.



In subsequent cycles the consultation trace can be considered to record all the spatial relationships between determinately located pairs of tokens present in the model.

Construction operations	Construction trace	Consultation trace (First Cycle)
1. Enter frog in model	Frog	Frog
2. Place duck in front of frog	Frog Duck	Frog (4) Duck (4)
3. Place mouse to right of frog	Frog Mouse Duck	Frog ( 3 ) Mouse ( 3 )
4. Place cat to right of frog	Frog ?Cat Mouse ?Cat Duck	
5. Place dog in front of cat	Frog ?Cat Mouse ?Cat Duck ?Dog ?Dog	

**Figure 6.3: An example of a construction trace and the first cycle of a consultation trace for the indeterminate description in figure 6.1.**

What is actually recorded in subsequent cycles will depend on how consultation is directed towards a particular goal or purpose. Thus the consultation trace in Figure 6.2 only records the ‘essential’ operations involved in checking the spatial mental model as it is being constructed. Subsequent cycles of the consultation trace for the determinate description in Figure 6.2 would record combinations of the possible spatial relationships between tokens in the spatial mental model. One plausible way for these to be recorded is for people to consult the spatial mental model one token at a time either to determine whether the spatial relations are correct or merely to refresh the spatial mental model. For the example in Figure 6.2 this might entail recording the spatial relations between frog and each of the remaining tokens in turn. This activity could then be repeated for each token in turn (though all possible relationships would be recorded at least once by the time the fourth token was reached). This sequence is attractive because it is relatively simple and because it is not inconsistent with the kinds of mental or sub-vocal rehearsal some people report. People would probably consult their spatial mental models in the order of spatial mental model construction either because early tokens have greater residual



activation from the construction trace or because they are rereading the description. Some people may not complete the whole sequence because of time pressures or because they are prepared to rely on working memory to recall later spatial relations. If this is the case this would lead to better long-term memory for early spatial relations, as has been observed. It should be stressed, however, that subsequent processing cycles are neither mandatory nor fixed. People are free to consult a spatial mental model in any way they wish.

The previous example of memory for a determinate spatial description describes the typical situation people are faced with when constructing a spatial mental model. Figure 6.3 above illustrates how the construction trace and the first cycle of a consultation trace might be recorded for a simple indeterminate description. When people are asked to read and understand a radically indeterminate spatial description both the construction and the consultation trace will differ. The main difference in the construction trace is that what is recorded is not the operation of adding a token to a spatial mental model but the operation of *attempting* to add a token to a spatial mental model. The main difference in the construction trace is that it is not possible to consult spatial relations between indeterminately located tokens and other tokens in the spatial mental model. These differences are based on the view that most people do not succeed in resolving the indeterminacy of the description, for instance by constructing only one of the possible models, by constructing two or more models or by annotating their model in some way (Johnson-Laird, 1983). While some people may adopt these strategies for coping with radical indeterminacy evidence from the experiments reported in Chapter 4 suggests that the majority of subjects do not or, if they do, are unsuccessful.

### Filling out the detail of the sketch

The examples presented earlier in this section are intended as the starting points of a theory of memory for spatial mental models which could account for the empirical evidence reported in this thesis and elsewhere. It has several features in its favour. Firstly, it can account for the absence or presence of an interaction between the determinacy of a description and the type of recognition performance, gist or verbatim, shown by subjects in several experiments. Where only minimal consultation of a spatial mental model is required people may not remember its structure (Payne, 1993). Where more extensive consultation of a spatial mental model is expected or required by subjects the crossover interaction predicted by Mani and Johnson-Laird (1982) should be and is observed. Secondly, by proposing that spatial mental model construction and consultation are ongoing, cyclic processes the sketch is consistent with evidence that the order and probability of recall are predicted by spatial mental model construction order.

It is important to note that differences in probability of recall and drawing order of items from determinate and indeterminate descriptions require that consultation processes also reflect the order of spatial mental model construction to some extent. This is incorporated in the sketch



by noting that people will tend to apply consultation processes in order to improve the long-term retention of the earlier tokens in spatial mental model construction. People will tend to rely more on working memory resources to retain information about later tokens in the construction order (Craik, 1970; Craik & Lockhart, 1972). Thus one of the most important characteristics of consultation processes is their role in elaborating the structure of a spatial mental model (Craik, 1979; Glenberg & Langston, 1992). This may seem counterintuitive, as subjects in the experiments reported here are only tested after intervening activity has been carried out. The explanation is that subjects are not told that they will be tested for delayed recall in any of the experiments reported here (the best example of this is the surprise recognition tests used in Chapter 4). People are used to relying on working memory resources for recall or recognition in everyday activities and it would be unusual for them to not make use of those resources in the experiments reported here.

What modifications need to be made to fill in the detail of this sketch? There are a number of features that could be altered in the formation of the construction trace. At present the sketch deals only with the simplest case where the construction process is terminated after a single cycle. In many cases people will refresh the construction trace by rereading the original description or by recalling and 're-running' the operations used to construct the spatial mental model. In some cases people may refresh the construction processes in a different order from that given in the original description (this is only likely if the entire description is available during the whole period of spatial mental model construction). It may also be necessary to argue that the first cycle of the construction trace results in a stronger memory trace than later cycles, to account for the good retention of construction processes when a description is heard rather than read. All these modifications are more or less minor changes to or extensions of the sketch which has been presented. There are also, however, more fundamental aspects of the sketch which have been left vague or which may be inadequate.

The first of these areas is the notion of activation or trace strength. At present the model copes with the notion of activation in three ways. Firstly the activation of arrays in long-term memory is partly indicated by the number of times a given array is recorded. Secondly, the number of times a given token is recorded in an array trace is also an indication of the level of activation of that token in long-term memory. Thirdly, tokens in the construction trace differ in activation according to their importance in the operation being recorded. New tokens have greater activation than existing tokens. Tokens acting as explicit reference points for new tokens have greater activation than those not explicitly mentioned in the relevant section of the description. This way of denoting activation or trace strength has considerable heuristic value, but should not be thought of as an accurate account of activation in long-term memory. A complete account of activation in long-term memory would almost certainly require a computational model of construction and consultation processes.



A second area which requires clarification is closely related to the issue of activation in long-term memory. This is the question of how the arrays which make up each trace are organized. One interpretation of the sketch is that each array is a separate long-term memory entity. This is problematic because there is considerable evidence that the different arrays can cue and be cued by each other (or by other information) in long-term memory. A preferable alternative is that the arrays are linked or associated within a larger long-term memory structure. Evidence for this view comes from experiments which demonstrate large fan effects only for information which can not be organized into a single coherent mental model (Radvansky et al., 1993; Radvansky & Zacks, 1991). One way of organizing different arrays into a single long-term memory representation would be to superimpose successive arrays onto a single 'meta-array'. Such a superpositional or distributed representation is often associated with connectionist systems (Rumelhart, 1989; Rumelhart & Norman, 1983; Rumelhart et al., 1986). Unlike distributed connectionist systems, however, the sketch presented here proposes that the construction and consultation trace explicitly preserve constituent structure (the tokens and the spatial relations between tokens). Distributed connectionist networks have been successful at modeling schemas, but it is not certain that they could model memory for spatial mental models (Fodor & Pylyshyn, 1988; Pinker & Prince, 1988; Rumelhart et al., 1986). An intermediate solution is that the arrays are organized temporally (from first to last). This solution is particularly good at explaining memory for order. The only significant problem with it being that construction and consultation processes do not proceed smoothly in a single cycle from first to last, but are likely to be repeated. This problem is less significant if it is assumed that identical arrays are superimposed on each other within the trace, resulting in increased activation for the original array. One way to explore this area may be to investigate long-term memory for the consultation processes involved in reasoning, because searching for counterexamples can involve constructing, consulting and discarding several mental models.

The third area of the sketch which is incomplete concerns the precise nature of the consultation processes. This aspect of the sketch is incomplete because specifying consultation processes operating on a spatial mental model requires a detailed analysis of the purpose for which the spatial mental model was constructed. To specify all the possible consultation processes requires a description of all the possible ways someone can consult a spatial mental model. There are, however, general properties of consultation processes which should be noted. People will tend to consult a spatial mental model to derive or generate new information rather than to derive information which is already available. This principle is already well known in the area of deductive reasoning where people avoid reaching redundant conclusions. If people are told "All acrobats are bakers" and "All bakers are chemists" people hardly ever conclude that "All acrobats are bakers and all bakers are chemists" even though this conclusion is perfectly valid (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). The implication is that when people have completed construction of a spatial mental model they will often actively



search for novel information (rather than information explicit in the premises) when they consult their representation.

The kinds of consultation processes operating on a spatial mental model are heavily task dependent. Some tasks will require people to derive a perceptual representation from the structure of a spatial mental model. Such a perceptual representation might take the form of a visual image (perceptual information could also be retrieved directly from long-term memory). Other consultation processes may require people to transform or annotate their spatial mental model in some way. In these cases consultation processes interact with construction processes. For example, two or three people in Experiment 4 reported constructing a spatial mental model and then transforming it into a linear string. Remembering the arrangement of objects simply required recalling the overall shape or annotating it. The determinate description above could thus be remembered as the string "Spoon, Fork, Plate, NEXT-LINE, Cup, Knife". Annotating a spatial mental model is also one way to represent an indeterminate description (though transforming an indeterminate spatial mental model into a string rather than array is not a productive strategy). To remember the indeterminate description above only two tokens ('Cat' and 'Mouse') need to be 'tagged' to indicate that their positions can legitimately be switched (it is also important to remember that if they are switched the token 'dog' remains in front of 'cat').<sup>3</sup> Annotated or augmented mental models are also postulated for certain kinds of reasoning and problem solving (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Newell, 1990).

One thing these consultation strategies share is that they are all consistent with the sketch of memory for spatial descriptions reported here. In order to transform a description into an annotated 'string' requires one to construct a representation of the spatial locations of each item in the original description. Similarly, before an indeterminate description can be represented as an annotated spatial mental model the mental modeler needs to notice that the description is indeterminate. The obvious way to do this is to attempt to construct an unannotated spatial mental model. Many consultation strategies also implicate construction processes in their application. It is convenient, though, to distinguish between construction processes (and perhaps the first cycle of consultation processes) and later consultation and construction processes which are less likely to follow a fixed pattern. This distinction between the construction and consultation of a spatial mental model has its parallel in the recent literature on mental models in reasoning. Valid deductive reasoning involves searching for counter-examples (a combination of consultation and construction processes), but many people only construct a single model of the

---

<sup>3</sup> This way of 'tagging' or annotating a spatial mental model is the strategy adopted by the author when constructing the materials used in Chapter 4. As such it has some introspective validity, but it is likely that it requires familiarity with the materials to be able to adopt the strategy readily (though at least one subject did report adopting a similar strategy at least some of the time).



premises which can lead to errors in reasoning (Oakhill & Johnson-Laird, 1985). Recent mental model research has shown that for other forms of reasoning searching for counter-examples is not an appropriate strategy (Johnson-Laird, 1993; Johnson-Laird, 1994). In solving meta-logical puzzles the strategies adopted by subjects are arguably the main focus of experimental research within the mental models framework (Byrne et al., 1992; Evans, 1990; Johnson-Laird & Byrne, 1990; Rips, 1990). This latter research has largely concentrated on working memory representation (e.g. one common method is to use protocol analysis). It would be interesting to determine whether evidence from studies of long-term memory representations would complement the findings obtained so far in this area.

### 6.3 Theoretical perspectives on memory for spatial mental models

The work presented here falls within the mental model 'framework' which proposes that people construct and manipulate integrated mental representations of real or imagined situations in order to understand and reason. According to this approach mental models provide a framework for a number of theories or sub-theories to tackle specific psychological issues such as anaphoric reference, deduction or induction (Garnham & Oakhill, 1992; Garnham & Oakhill, 1994). This thesis can be considered as contributing to a sub-theory of how people construct and manipulate spatial mental models, but the implications of this research are not confined only to mental models. The interpretation of, and the assumptions behind, this research share common ground with other theories of language comprehension and reasoning and with more general strands of research in cognitive psychology. These include work on imagery, depth of processing and implicit memory. A selection of these connections form the basis of the brief discussion in this section.

The interpretation of the empirical research reported here is broadly compatible with existing accounts of mental or situation models in reasoning and discourse comprehension (Garnham & Oakhill, 1994; Glenberg & Langston, 1992; Glenberg et al., 1987; Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Tversky, 1991; van Dijk & Kintsch, 1983). It is also more generally related to approaches to reasoning and problem solving which propose integrated representations of the task (Evans et al., 1993; Galotti, 1989; Newell, 1990; Potts, 1972). Theories of relational reasoning, for example with three-term series problems, using integrated representations of the premises have a long history in psychology and have been particularly influential in the formation of theories of reasoning based on mental models (see section 1.2.4). In the area of psycholinguistics this mental model approach shares many features with constructivist approaches to language comprehension (Bransford et al., 1972; Bransford & Franks, 1971; Schank, 1982; Schank & Abelson, 1977). One recent formulation of the constructivist approach to language comprehension deserves particular attention. The *structure building framework* of Gernsbacher (1990; 1991) proposes three processes which make up comprehension.



The first process lays a foundation for a coherent mental structure of the discourse. The second process maps new information onto an existing structure. The third structure shifts to a new structure if the construction of single coherent structure is not possible. Each of these processes is controlled by one of two mechanisms: *enhancement*, which increases activation, and *suppression*, which reduces activation. Evidence for the framework comes from studies showing that items mentioned first in discourse are processed longer 'on-line' and remembered better 'off-line'. The similarities between the mental models and structure building frameworks are not surprising (both emerged from the same constructivist psycholinguistic tradition). What is interesting is that the structure building framework makes similar predictions concerning order and probability of recall as the thesis presented here. Gernsbacher, however, predicts that there is always an advantage for the first item mentioned, whereas the account predicted here proposes that in some instances the second item mentioned is the first to be entered into a spatial mental model (or into a 'mental structure'). This interpretation, though, is implicit in the framework proposed by Gernsbacher provided the item entered first into the structure rather than the item mentioned first in the description is stressed. The work by Gernsbacher offers important insights into concepts such as activation which have been neglected by the mainstream mental models tradition and should be seen as complementary rather than in opposition to the mental models framework.

Theories of imagery can also be related to the mental models approach. These theories of imagery fall roughly into two kinds both of which deal with the construction of special-purpose representations in working memory. The first kind of imagery theory is exemplified by the work of Kosslyn (Kosslyn, 1980; 1981; 1991). Kosslyn describes a computational model of how images are constructed and manipulated in a visuo-spatial buffer. Like the account presented in this thesis Kosslyn makes no claims that images are represented in a special format in long-term memory. Kosslyn does distinguish between the structural and visual aspects of imagery in his computational model, but the two interact in order to account for classical imagery effects (Kosslyn, 1980). For instance, the Kosslyn model does not distinguish between perceptual information retrieved from memory and perceptual information generated from a structural representation. In many cases visual and structural information do interact, but it is also possible to have a spatial representation which is not visual (Conway, 1983). For example, an auditory image can also convey spatial information in the absence of visual imagery. In most cases, though, the kinds of 'spatial' images proposed by Kosslyn are very similar to the spatial mental models described here, sequentially constructed and manipulated special-purpose representations in working memory (Kirby & Kosslyn, 1990; Kosslyn et al., 1988).

The second kind of imagery theory related to the work presented here is that which describes how working memory encoding or elaboration processes could account for the memorability of pictures or concrete words. Imagery which links or relates two or more items together is shown to lead to better retention than imagery which stresses the visual features of



individual words or objects (Bower, 1970; Hunt & Einstein, 1980; Marschark, 1985; Marschark & Surian, 1989; Morris & Stevens, 1974; Morris and Reid, 1975). In the work presented here it is elaboration (a combination of the amount and the relational nature of the encoding processes) which is thought to determine the memorability of items within a single spatial mental model. This differs from depth of processing and other elaboration accounts in that it is confined to the circumstances where people construct an integrated representation of a situation. No strong claims are made to suggest that all imagery or elaboration accounts can be reduced to a single explanation based on the construction of a mental model, though in many everyday situations people do appear to benefit from the construction of such integrated representations (Cohen, 1989; Evans et al., 1993; Garnham & Oakhill, 1994; Glenberg & Langston, 1992; Johnson-Laird, 1994).

While memory for spatial mental models does appear to be related to integrative, elaborative processing, there is also evidence, particularly in memory for construction processes, that transfer-appropriate processing is important. Unlike the original transfer-appropriate processing hypothesis the view of memory for construction processes favoured here is explicitly representational and assumes that arrays and tokens in the construction trace can potentially be accessed individually as well as by re-activating or re-running the whole operation (Bransford et al., 1979; Craik, 1979; Morris et al., 1977). If this were not the case, greater burden would be placed on consultation processes to account for memory for the order of spatial mental model construction. Construction processes are more constrained than consultation processes and so it is more appropriate for some of this burden to be met by memory for a construction trace. Recent research on implicit memory has revived transfer-appropriate processing explanations (Roediger, 1990; Weldon & Roediger, 1987). It is worth considering that memory for construction processes may reflect implicit memory involvement in an explicit memory task. It is certainly the case that people have explicit recall of spatial relations, but there is little evidence in the experiments reported in this thesis that people are consciously aware of recalling the process of constructing a spatial mental model.

## 6.4 Methodological issues

In the course of this thesis a number of methodological issues have been addressed. In the sections below the key components of these issues will be summarized and discussed. For discussion of methodological points relating to individual experiments the reader is referred to the empirical work in Chapters 3 to 5.



### **6.4.1 Distinguishing between different forms of mental representation**

This thesis has been written within a psychological tradition which proposes that it is possible to distinguish empirically between different forms of mental representation. In one sense this is less a methodological point than a philosophical one. The most fundamental distinction, however, is not between different forms of representation as such, but between different levels of representation (Johnson-Laird, 1983; Palmer, 1978). It has been argued that it is possible to distinguish between different forms of high-level representation using methods available to cognitive psychology, at least in working memory. These representations are defined by the kinds of high-level processes that they support and the kind of processes that are available to act upon them, not by some primitive 'machine code' (Johnson-Laird, 1983; Palmer, 1978). The question of what kind of primitive or low-level representations are involved in cognition is one that is outside the realms of contemporary cognitive psychology. Answering this question requires a solution to the mind-body problem: the central problem of cognitive science (Bechtel, 1988; Lycan, 1990). The chief defence of the view presented here does not rest on empirical findings per se, but on the ability of research which adopts this viewpoint, such as this thesis, to be productive and to reach internally consistent findings.

### **6.4.2 Distinguishing between working memory and long-term memory representations**

Payne (1993) has already provided an eloquent argument on the importance of distinguishing between the representation of spatial mental models in working memory and long-term memory. These points are addressed earlier in the thesis, but the gist of the argument is worth repeating. Most studies of mental models do not force people to remember the mental models they have constructed. The majority require people to construct a mental model and, for example, report a conclusion reached by consulting it (Johnson-Laird & Byrne, 1991). In some cases we have good reason to suppose that long-term memory is also implicated, for example because the experiment involves reading a long text or consulting a spatial mental model over a long period of time (Bower & Morrow, 1990; Kintsch, 1986; Tversky, 1991). Nevertheless, in almost every case, there is nothing to prevent people from relying heavily (and in some cases totally) on working memory to perform the task. The solution proposed by Payne (1993), and adopted for every experiment reported here, is to prevent people using temporary storage capacity in working memory by introducing an intervening task. Two types of intervening tasks were used here, usually in conjunction. The first type involves presenting a task requiring a second spatial mental model to be constructed in working memory. This is particularly useful because mental model construction is held by most authors to be very working memory intensive (Johnson-Laird, 1983; Tardieu et al., 1992). The second type of task involves recalling or recognizing a spatial description presented earlier. Both types of tasks can introduce considerable delays between presentation and retrieval



and carry a risk of confusion or interference between descriptions or spatial mental models. Delays or interference, however, will in general make it less rather than more likely to find evidence of memory for any aspect of a spatial description. These experiments can only be interpreted in terms of long-term memory or in terms of the retrieval from long-term memory into working memory of information derived from the original spatial descriptions.

### 6.4.3 Ecological validity

Ecological validity is an important and sometimes controversial topic in modern cognitive psychology and in memory research in particular. Ecologically valid or everyday memory studies produce findings which are readily related to real world applications or situations and which still emerge when many incidental factors are taken into consideration (Cohen, 1989; Cohen et al., 1993). Laboratory studies may lack these advantages, but are often more readily generalized to novel situations and controls can be applied to exclude incidental factors which are not of direct interest (Banaji & Crowder, 1989). In the empirical work presented here ecological validity is of concern in the nature of the materials presented to the subjects and in the nature of the tasks employed. Some of the experiments have employed relatively rich and meaningful stimuli such as stories with a strong narrative component. Other stimuli have employed relatively simple and meaningless stimuli composed of sets of semantically related items and spatial relations connecting them. Similar findings have been obtained with both kinds of stimuli, though the simpler materials have permitted more precise control of the content and order of presentation (for example, compare experiments 3 and 5).

The similarity in the conclusions from experiments using a range of different materials helps guard against problems inherent in both everyday and traditional laboratory studies. The tasks subjects perform are ecologically valid only in a limited sense. They all involve understanding, constructing and consulting a spatial representation derived from a verbal description. This is something that people do frequently and relatively effortlessly. If it were not the case people would not be able to provide directions, interpret road signs and so on. Many potential difficulties in navigation and spatial cognition are mitigated by redundancy in the way we communicate and record spatial information and from the many sources which surround us (road signs are usually repeated, directions are often given verbally *and* as maps, or attempt to combine canonical and relative direction information). In this sense the tasks employed in this thesis are not always high in ecological validity. Nor were the retrieval situations very high in ecological validity (though most subjects preferred drawing spatial layouts to any of the other testing situations). Overall then, these experiments could be regarded as fairly traditional laboratory studies. Where this research departs from such traditional experimental methods is in a deliberate attempt to demonstrate memory for spatial mental models over a range of materials, tasks and retention intervals which share a common goal. This goal is to explore the



way people construct representations of spatial information and how these representations influence long-term memory.

#### 6.4.4 Strategies and individual differences

Individual differences are often held to be problematic for many models of cognitive processes to deal with. This form of analysis, though, falsely assumes that because a model does not explicitly address individual differences it cannot explain any individual differences that do arise. Where models in cognitive psychology do seek to explain individual differences they can be very successful. For example individual differences in working memory can be very powerful predictors of performance in reasoning, problem solving and some areas of language comprehension (Daneman & Carpenter, 1980; Hunt, 1987; Just & Carpenter, 1992; Kyllonen & Christal, 1990). Individual differences at the level of cognitive architecture (like working memory) are where cognitive psychology can be particularly productive. One area where individual differences have been particularly controversial has been in imagery research. Experience of visual or other imagery varies immensely from person to person, perhaps more than any other aspect of conscious experience - the question of imageless thought being one of the oldest controversies in psychology (Boring, 1950; Dellarosa, 1988; Hilgard, Leary, & McGuire, 1991). The problem of imagery experience for cognitive psychology is cognitive psychologists try to discover general cognitive processes rather than faculties that are possessed by some individuals but not others. The mental models framework avoids this problem by proposing that the ability to construct mental models is universal and that the subjective experience of visual or other imagery is distinct from this. Thus differences in imagery do not influence deduction with spatial mental models, even though some people experience visual imagery during deductive reasoning (Evans et al., 1993).

Until recently individual differences have not been a major focus of the original theory of mental models (Johnson-Laird, 1983). Mental models, it is argued, predict individual differences because of working memory limitations or because they permit people to construct mental models which are influenced by real world knowledge. For instance, some people may reason differently by adding to the information contained in the premises. In extreme cases real-world knowledge is used to reject valid conclusions generated from a mental model (Oakhill & Garnham, 1993; Oakhill et al., 1990; Oakhill et al., 1989). Expert-novice differences are also often characterized by a shift from reasoning with superficial representations of a problem towards reasoning with a detailed, domain-specific mental model (Chi, Feltovich, & Glaser, 1981; Kintsch, 1994; Tardieu et al., 1992). Arguably, even novices will construct a mental model, but it will rarely be domain-specific and will often be too abstract to reason accurately with. The account of memory for spatial mental models offered here goes beyond standard mental model accounts by emphasising the importance of different kinds of consultation process on what is



remembered. These consultation processes reflect more or less deliberate strategies adopted for a specific task or class of problems. This emphasis on strategies adopted in accessing information from spatial mental models parallels a growing interest in the strategies adopted in deductive and inductive reasoning within the mental models framework (this was discussed in more detail at the end of section 6.2.2). One important consideration is that these strategies are not ad hoc inventions, because they are all derived from a common assumption - that comprehension and reasoning require a representation of the structure of a situation. If people are able to construct mental models it follows that they are able to reach valid deductive conclusions *if* they search for counter-examples (Johnson-Laird, 1983). If people are able to construct mental models it follows that they can make spatial inferences simply by reading information off from those models (Byrne & Johnson-Laird, 1989). These strategies are only ad hoc if they are considered in isolation from the ability to construct and consult a mental model.

## 6.5 Representational issues

In this section three representational issues of central importance to this thesis and to the mental models framework are discussed. The first issue has been addressed at several points in this thesis, (notably in Chapters 1 and 3) and concerns the distinction between mental models and images. The second issue has been addressed in the interpretation of several of the experiments in the thesis and concerns the distinction between schemas and mental models. The third issue is the main focus of the thesis - the question of how spatial mental models are or are not represented in long-term memory.

### 6.5.1 The relationship between images and mental models

One objection to the idea that people construct and manipulate mental models has been that mental models can not be distinguished from images. According to Johnson-Laird (1983) an image is a "view" of a mental model. This has been interpreted in this thesis to mean that an image corresponds to the perceptual rather than the structural component of a situation. Not all mental models are visual or spatial in nature. Reasoning often involves constructing mental models which represent negation (Johnson-Laird & Byrne, 1991). Mental models can also be constructed to represent procedures (Glenberg & Langston, 1992). In defence of this position the finding that it is possible to experience rich visual or other imagery which does not possess structure can be cited. This has been most clearly demonstrated in neurophysiology and neuropsychology (Farah, 1988; Farah et al., 1988). Memory research suggests that complex images derived from autobiographical experience do not need to be constructed sequentially like complex images derived from semantic memory (Conway, 1983). One explanation of this is that perceptual information about the episode can be retrieved directly, while structure has to be generated



'piece by piece'. Research by Taylor and Tversky (1992a) has shown that people can construct "perspective-free" representations of spatial situations, and their observation that drawing an environment tends to reflect the order of spatial mental model construction (rather than a visual image) has been confirmed here. This distinction between images and spatial mental models can explain why individual differences in visual imagery have a very poor record in accounting for differences in reasoning ability (see section 6.4.4 above). Visual imagery ability can only influence tasks which require the generation, retrieval or manipulation of perceptual rather than structural information. This is very different from arguments which claim that imagery experience is epiphenomenal (Pylyshyn, 1973; 1984). Imagery experience may or not play a causal role in cognitive processes according to whether those processes can benefit from understanding the perceptual characteristics of a situation.

### 6.5.2 Schemas and mental models in long-term memory

A different kind of objection against mental models suggests that mental models are simply an alternative term for schemas. This objection is based on the assumption that mental models, like schemas, are representations in long-term memory, an assumption which for the mental models described by Johnson-Laird at least, is false. A more sophisticated version of this objection is made by Brewer (1987). Brewer suggests that mental models are instantiations of long-term memory schemas in working memory.<sup>4</sup> Thus a mental model is constructed by filling in slots, assigning defaults and so on.

This view is partially accurate, in that this is one way in which a mental model can be constructed, but it is not the only way. The experiments presented in this thesis provide one of the clearest demonstrations of this. Memory for mental models and memory for spatial mental models is episodic and specific. Spatial relations between entities rarely conform to knowledge stored in long-term memory. This is not to say that you cannot have a schema for a typical English village or a typical university campus (see Experiments 7 and 8). However, specific spatial relations contained in the schema will be next to useless. For example, draw a map of a typical university campus and then try and use that map to navigate round a university campus you have never been to. A great deal of non-spatial information on the map would be correct (e.g. that there is a union building or halls of residence). Nearly all of the spatial information on the map will be incorrect (important buildings may tend to be more central, but it is extremely unlikely that many buildings will be in appropriate spatial relationships to each other). Schemas can not explain memory for highly specific situations. In addition the episodic nature of memory for

---

<sup>4</sup> Brewer does suggest labelling these kinds of mental models *episodic mental models*. Nevertheless it is misleading to suggest they are merely instantiations of schemas.



construction processes is inconsistent with a schema interpretation of memory for spatial mental models. Mental models are episodic representations of specific situations which can be constructed from discourse, perception or organized knowledge in long-term memory. Mental models may interact with schemas, be constructed from schemas and even form the basis of new schemas, but this not sufficient to make them schemas or instantiations of schemas.

### 6.5.3 The format of spatial mental models in long-term memory

Earlier in this chapter the idea of a simple 'copy' theory of memory for spatial mental models was rejected. Instead a theory based on memory for the processes of constructing and consulting a spatial mental model was favoured, and a sketch of such a theory was offered. This approach has similarities with theories of imagery or mental models which propose that long-term memory is unitary and propositional (Glenberg & Langston, 1992; Kosslyn, 1980; Payne, 1993). The construction and consultation traces proposed in section 6.2.2 are not, however, described as sets of propositions, but as sets of arrays. These arrays can be considered in several different ways. One way to consider retrieval from these arrays is that they in some sense reinstate conditions in working memory. This view is akin to transfer-appropriate processing (Morris et al., 1977; Weldon & Roediger, 1987). A second way to consider them is as high-level representations similar to structural descriptions (Johnson-Laird, 1989; Tversky, 1991). Each viewpoint captures the main idea behind the construction and consultation trace. This is that the spatial or structural information *being processed* is preserved in the trace. In the former case it is preserved because retrieval somehow reinstates the processes acting on the structure of the original spatial mental model at some point in its construction or consultation. How exactly this kind of reinstatement is achieved is the major problem that faces transfer-appropriate processing accounts (Bransford et al., 1979; Craik, 1979). In the latter case the succession of arrays or 'structural descriptions' which make up the trace are laid down in long-term memory as processing is carried out (Palmer, 1978; Rumelhart & Norman, 1983).

One of the most difficult questions to answer is whether the traces encoded in long-term memory share the same format as those in working memory. An affirmative answer to this question requires that the processes acting upon the long-term memory traces do so in the same way as the special-purpose processes acting upon spatial mental models in working memory (Johnson-Laird, 1983; Palmer, 1978). The tentative response offered here is that retrieval involves reinstatement of processes or structures into working memory and hence mental modeling per se is confined to working memory. This would be supported if accessing information from the trace increases with the complexity or size of the array (this might not be the case if arrays can be acted on directly by procedures in long-term memory). In either case working memory mental models differ in *content* from the consultation and especially the construction trace. A working memory mental model does not necessarily represent information about the order and process of



its construction, unlike the long-term memory traces described here. In some cases, where consultation processes are minimal, the long-term memory traces will only encode part of the structural information from a working memory mental model. Where consultation processes are extensive the long-term memory trace will approximate to the structure of the completed spatial mental model.<sup>5</sup> In these cases the long-term memory traces can act as an analog of the completed spatial mental model in working memory, but they are not a copy of it.

## 6.6 Implications for future research

It is dangerous to assume that representations in working memory are identical to or copies of representations in long-term memory. Memory for spatial mental models provides a good example of this. In everyday activity people switch between long-term memory and working memory almost effortlessly. Working memory is capacity limited, but long-term memory is for all practical purposes unlimited. This ability to extend working memory by accessing organized structures in long-term memory is a fundamental aspect of human expertise (Chase & Simon, 1973; Chi et al., 1981). This presents problems for cognitive psychologists. It is not always easy to tell where long-term memory begins and working memory ends. Long-term memory reflects ongoing, dynamic processing in working memory not just a particular state at a particular point in time. One of the challenges for cognitive science is to provide ‘unified’ accounts of cognitive processes (Newell, 1990; 1992). One important step along the way is a better understanding of how working memory and long-term memory interact (Baddeley, 1992). Within the mental models framework this requires a better understanding of the time course of mental model construction and consultation processes. The sketch of memory for spatial mental models described in section 6.2.2 is one possible starting point for this. Consultation processes are driven by the conscious strategies people adopt for particular tasks and situations. A better understanding of these processes thus requires an account of how strategies are developed for different kinds of situations. Lastly, the account outlined here for spatial mental models may form the basis of a general account of mental modeling processes. This in turn requires an understanding of the extent to which spatial mental modeling draws on processes or resources which are specific to spatial cognition. Recent research has already begun to address both strategy use (Byrne et al., 1992; Johnson-Laird, 1994) and the nature of working memory resources involved in constructing and manipulating a mental model (de Vooght & Vandierendonck, 1993; Gilhooly et al., 1993; Toms et al., 1993).

---

<sup>5</sup> This assumes that spatial mental model construction culminates in a final, complete, static representation as in the examples in section 6.2.2.



Several of the more speculative aspects of the conclusions reached in this thesis suggest further lines of research. The differences in memory for spatial information when verbal descriptions are read and when they are heard could be investigated further. If it is true that people hearing a verbal description are less likely to deviate from the predicted mental model construction order then this should be reflected in a stronger relationship between spatial mental model construction order and order or probability of recall. It has also been proposed that the construction trace is procedural rather than declarative in nature. People rarely, if ever, explicitly mention remembering the process of spatial mental model construction. This in turn may be related to the implicit nature of the measures used to test the episodic construction trace hypothesis. Payne (personal communication) has noted the similarity between his hypothesis and recent implicit memory research. Implicit memory methodologies may provide a useful framework for investigating differences between the construction trace and consultation trace.

One important belief which has driven this research is that spatial mental modeling is more similar than dissimilar to mental modeling in other domains. There is already some evidence that similar findings have been obtained with the comprehension of picture stories (Gernsbacher, 1990; 1991), it would be useful to extend the sketch presented here to account for long-term memory on other spatial tasks (such as spatial reasoning or real-world navigation) or in other domains (such as understanding temporal or causal relationships). Lastly, several areas of the sketch presented here require further investigation. The most important of these is probably the relationship between the construction processes and the consultation processes. Two ways to investigate the relationship between the construction and consultation trace suggest themselves. The first is to systematically vary demands placed on subjects during the initial learning or orienting task. Low task demands should make it more likely that people will remember only the process of spatial mental model construction. High task demands should make it more likely that people will extensively consult a spatial mental model and so remember more of the structure of the spatial mental model. The second way is to look at the effects of different retention intervals on recognition or recall. The evidence presented in this thesis suggests that the structure of a spatial mental model, and hence a consultation trace, can be retained even at very long retention intervals.



## References

Albrecht, J.E. & O'Brien, E.J. (1993). Updating a mental model: maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, pp. 1061-1070.

Altham, P.M. (1973). A non-parametric measure of signal discriminability. *British Journal of Mathematical and Statistical Psychology*, 26(1), pp. 1-12.

Anderson, J.R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 85, pp. 249-277.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J.R. (1984). Spreading activation. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honour of Gordon Bower*, pp. 61-90. San Francisco, CA: W.H. Freeman.

Anderson, J.R. & Bower, G.H. (1973). *Human associative memory*. Washington, D.C.: Hemisphere Press.

Anderson, J.R. & Reder, L.M. (1979). An elaborative processing explanation of depth of processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 385-403. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, R.C. & Pichert, J.W. (1978). Recall of previously unrecalable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 17, pp. 1-12.

Anderson, T., Tolmie, A., Howe, C., Mayes, T. & Mackenzie, M. (1992). Mental models of motion. In Y. Rogers, A. Rutherford & P.A. Bibby (Eds.), *Models in the mind: theory, perspective and application*, pp. 57-71. London: Academic Press.

Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.



- Baddeley, A.D. (1988). But what the hell is it for? In M.M. Gruneberg, P.E. Morris & R.N. Sykes (Eds.), *Practical aspects of memory: current research and issues, Volume 1: memory in everyday life*. pp. 3-18. Chichester: John Wiley.
- Baddeley, A.D. (1990). *Human memory: theory and practice*. Hove: Lawrence Erlbaum Associates.
- Baddeley, A.D. (1992). Working memory: the interface between memory and cognition. *Journal of Cognitive Neuroscience*, 4, pp. 281-288.
- Baddeley, A.D. & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 8. pp. 47-89. San Diego: Academic Press.
- Baddeley, A.D., Logie, R., Nimmo-Smith, I. & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, pp. 119-131.
- Baddeley, A.D. & Wilson, B. (1988). Comprehension and working memory: a single case neuropsychological study. *Journal of Memory and Language*, 27, pp. 479-498.
- Baguley, T.S. (1992). Investigating spatial representations through learning and problem solving with maps and descriptive texts. Paper presented at the *Second Interdisciplinary Workshop on Mental Models: Mental Models and Everyday Activities*. March 23-25, Robinson College, Cambridge.
- Bahrnick, H.P. (1970). Two-phase model for prompted recall. *Psychological Review*, 77, pp. 215-222.
- Bahrnick, H.P. (1983). The cognitive map of a city: 50 years of learning and memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 17. pp. 125-163. San Diego: Academic Press.
- Bahrnick, H.P. (1984). Semantic memory content in the permastore: fifty years of memory for Spanish learned in high school. *Journal of Experimental Psychology: General*, 104, pp. 54-75.
- Banaji, M.R. & Crowder, R.G. (1989). The bankruptcy of everyday memory. *American Psychologist*, 44, pp. 1185-1193.
- Baron, J. (1988). *Thinking and deciding*. Cambridge: Cambridge University Press.
- Bartlett, F.C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bechtel, W. (1988). *Philosophy of mind: an overview for cognitive science*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Black, J.B. (1984). Understanding and remembering stories. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honour of Gordon Bower*, pp. 235-255. San Francisco: W.H. Freeman.
- Boden, M. (1988). *Computer models of mind: computational approaches in theoretical psychology*. Cambridge: Cambridge University Press.



Boring, E.G. (1950). *A history of experimental psychology* (2nd Ed. ed.). New York: Appleton Century.

Bower, G.H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behaviour*, 9, pp. 529-533.

Bower, G.H. (1989). Mental models in text understanding. In A.F. Bennett & K.M. McKonkey (Eds.), *Cognition in individual and social contexts*, pp. 129-144. Amsterdam: Elsevier.

Bower, G.H. & Morrow, D.G. (1990). Mental models in narrative comprehension. *Science*, 247, pp. 44-48.

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, pp. 1-21.

Bransford, J.D., Barclay, J.R. & Franks, J.J. (1972). Sentence memory: a constructive versus interpretive approach. *Cognitive Psychology*, 3, pp. 193-209.

Bransford, J.D. & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, pp. 331-350.

Bransford, J.D., Franks, J.J., Morris, C.D. & Stein, B.S. (1979). Some general constraints on learning and memory research. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 331-354. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brewer, W.F. (1987). Schemas versus mental models in human memory. In P. Morris (Eds.), *Modelling cognition*, pp. 187-197. Chichester: John Wiley.

Brewer, W.F. & Pani, J.R. (1983). The structure of human memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 17, pp. 1-38. San Diego: Academic Press.

Brewer, W.F. & Treyens, J.C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13, pp. 207-230.

Broadbent, D.E. (1985). A question of levels: comment on McClelland and Rumelhart. *Journal of Experimental Psychology: General*, 114, pp. 189-190.

Brown, R. & McNeill, D. (1966). The "tip-of-the-tongue" phenomenon. *Journal of Verbal Learning and Verbal Behaviour*, 5, pp. 325-337.

Bruce, V.B. & Green, P.R. (1985). *Visual perception: physiology, psychology and ecology*. London: Lawrence Erlbaum Associates.

Byrne, R.M.J., Handley, S.J. & Johnson-Laird, P.N. (1992). Advances in the psychology of reasoning: meta-deduction. In M.T.G. Keane & K.J. Gilhooly (Eds.), *Advances in the psychology of thinking*, pp. 127-145. London: Harvester.

Byrne, R.M.J. & Johnson-Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language*, 28, pp. 564-575.



- Byrne, R.M.J. & Johnson-Laird, P.N. (1990). Remembering conclusions we have inferred: what biases reveal. In J.-P. Caverni, J.-M. Fabre & M. Gonzalez (Eds.), *Cognitive biases*, pp. 109-120. Amsterdam: North-Holland.
- Chase, W.G. (1983). Spatial representations of taxi drivers. In D.R. Rogers & J.A. Sloboda (Eds.), *Acquisition of symbolic skills*, pp. 391-405. New York: Plenum.
- Chase, W.G. & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, 4, pp. 55-81.
- Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, pp. 121-152.
- Chomsky, N. (1980). Rules and representations. *Behavioural and Brain Sciences*, 3, pp. 1-61.
- Clark, H.H. (1969). Linguistic processes in deductive reasoning. *Psychological Review*, 76, pp. 387-404.
- Clark, H.H. (1973). Space, time, semantics and the child. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language*, pp. 27-63. New York: Academic Press.
- Cohen, G. (1983). *The psychology of cognition* (2nd. ed.). London: Academic Press.
- Cohen, G. (1989). *Memory in the real world*. Hove: Lawrence Erlbaum Associates.
- Cohen, G. (1990). Recognition and retrieval of proper names: age differences in the fan effect. *European Journal of Cognitive Psychology*, 2, pp. 193-204.
- Cohen, G., Conway, M. & Stanhope, N. (1992). *Very long term retention of knowledge*. (HCRL Technical Report 87). The Open University.
- Cohen, G. & Faulkner, D. (1986). Memory for proper names age differences in retrieval. *British Journal of Developmental Psychology*, 4, pp. 187-197.
- Cohen, G., Kiss, G. & Le Voi, M. (1993). *Memory: current issues* (2nd ed.). Milton Keynes: Open University Press.
- Conway, M.A. (1983) *Content and organizational differences between autobiographical and semantic memories*. Unpublished Ph.D. Thesis, Open University.
- Conway, M.A. (1990). *Autobiographical memory: an introduction*. Milton Keynes: Open University Press.
- Cooper, L.A. (1976). Demonstration of a mental analog of external rotation. *Perception and Psychophysics*, 19, pp. 296-302.
- Craik, F.I.M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behaviour*, 9, pp. 143-148.
- Craik, F.I.M. (1979). Levels of processing: overview and closing comments. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 447-461. Hillsdale, NJ: Lawrence Erlbaum Associates.



Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, 11, pp. 671-684.

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, pp. 268-294.

Craik, K.J.W. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19, pp. 450-466.

de Vega, M. (1992). *Mental models of characters in spatial scenarios*. Paper presented at the Fifth European Conference on Cognitive Psychology, September 12-16, Paris.

de Vooght, G. & Vandierendonck, A. (1993). *On the spatial representation of time concepts in reasoning tasks*. (Reports of the Department of General Psychology No. 10). University of Ghent.

Dellarosa, D. (1988). A history of thinking. In R.J. Sternberg & E.E. Smith (Eds.), *The psychology of human thought*, pp. 1-18. Cambridge: Cambridge University Press.

Denhière, G. & Denis, M. (1988). The processing of texts describing spatial configurations. In H. Mandl & J.R. Levin (Eds.), *Knowledge acquisition from text and pictures*, pp. 249-261. Amsterdam: Elsevier.

Denis, M. (1991). *Image and cognition*. London: Harvester.

Ehrlich, K. & Johnson-Laird, P.N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behaviour*, 21, pp. 296-306.

Ehrlich, M.-F. & Tardieu, H. (1992). *Conceptual models in text comprehension*. Paper presented at the Fifth European Conference on Cognitive Psychology, September 12-16, Paris.

Ellis, N. (1991). Word meaning and the links between the verbal system and, modalities of perception and imagery. Or, in verbal memory the eyes see vividly, but ears only faintly hear, fingers barely feel and the nose doesn't know. In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition*, pp. 313-329. Amsterdam: North-Holland.

Evans, J.S.B.T. (1990). Reasoning with knights and knaves: a discussion of Rips. *Cognition*, 36, pp. 85-90.

Evans, J.S.B.T. (1991). Theories of human reasoning: the fragmented state of the art. *Theory and psychology*, 1, pp. 83-105.

Evans, J.S.B.T. (1993). The cognitive psychology of reasoning: an introduction. *Quarterly Journal of Experimental Psychology*, 46A, pp. 561-567.

Evans, J.S.B.T., Newstead, S.E. & Byrne, R.M.J. (1993). *Human reasoning: the psychology of deduction*. Hove: Lawrence Erlbaum Associates.

Eysenck, M.W. (1979). Depth, elaboration, and distinctiveness. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 89-118. Hillsdale, NJ: Lawrence Erlbaum Associates.



- Eysenck, M.W. (1984). *A handbook of cognitive psychology*. London: Lawrence Erlbaum Associates.
- Eysenck, M.W. & Eysenck, M.C. (1979). Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Journal of Experimental Psychology: Human Perception and Performance*, 5, pp. 472-484.
- Eysenck, M.W. & Keane, M.T. (1990). *Cognitive psychology: a student's handbook*. Hove: Lawrence Erlbaum Associates.
- Farah, M.J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, pp. 307-317.
- Farah, M.J., Hammond, K.M., Levine, D.N. & Calvanio, R. (1988). Visual and spatial mental imagery: dissociable systems of representation. *Cognitive Psychology*, 20, pp. 439-462.
- Farah, M.J., Peronnet, F., Gonon, M.A. & Giard, M.H. (1988). Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *Journal of Experimental Psychology: General*, 117, pp. 248-257.
- Fletcher, C.R. (1992). Assessing recognition memory for surface forms in discourse: a methodological note. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, pp. 199-203.
- Fletcher, C.R. & Chrysler, S.T. (1990). Surface forms, textbases and situation models: recognition memory for three types of textual information. *Discourse Processes*, 13, pp. 175-190.
- Fodor, J.A. (1975). *The language of thought*. New York: Crowell.
- Fodor, J.A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioural and Brain Sciences*, 3(1), pp. 63-110.
- Fodor, J.A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press/Bradford Books.
- Fodor, J.A. (1987a). *Psychosemantics*. Cambridge, MA: MIT Press/Bradford Books.
- Fodor, J.A. (1987b). Modules, frames, fridgeons, sleeping dogs and the music of the spheres. In J.L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding*, pp. 25-36. Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J.A. & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28, pp. 3-71.
- Franklin, N., Tversky, B. & Coon, V. (1992). Switching points of view in spatial mental models acquired from text. *Memory and Cognition*, 20, pp. 507-518.
- Freebody, P. & Anderson, R.C. (1986). Serial position and rated importance in the recall of text. *Discourse Processes*, 9, pp. 31-36.
- Galotti, K.M. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin*, 105, pp. 331-351.



Gardner, H. (1987). *The mind's new science: a history of the cognitive revolution* (2nd ed.). New York, NY: Basic Books.

Garnham, A. (1981). Mental models as representations of text. *Memory and Cognition*, 9, pp. 560-565.

Garnham, A. (1987). *Mental models as representations of discourse and text*. Chichester: Ellis Horwood.

Garnham, A. & Oakhill, J. (1992). Discourse processing and text representation from a "mental models" perspective. *Language and Cognitive Processes*, 7, pp. 193-204.

Garnham, A. & Oakhill, J. (1994). *Thinking and reasoning*. Oxford: Blackwell.

Garnham, A., Oakhill, J.V. & Johnson-Laird, P.N. (1982). Referential continuity and the coherence of discourse. *Cognition*, 11, pp. 29-46.

Gathercole, S.E. & Baddeley, A.D. (1993). *Working memory and language*. Hove: Lawrence Erlbaum Associates.

Gentner, D. & Collins, A. (1990). Multiple models of evaporation processes. In D.S. Weld & J. De Kleer (Eds.), *Readings in qualitative reasoning about physical systems*, pp. San Mateo, CA: Morgan Kaufmann.

Gentner, D. & Gentner, D.R. (1983). Flowing waters or teeming crowds: mental models of electricity. In D. Gentner & A.L. Stevens (Eds.), *Mental models*, pp. 99-129. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D. & Stevens, A.L. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gernsbacher, M.A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, 17, pp. 324-363.

Gernsbacher, M.A. (1990). *Language comprehension as structure building*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gernsbacher, M.A. (1991). Cognitive processes and mechanisms in language comprehension: the structure building framework. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 27, pp. 217-263. San Diego: Academic Press.

Gernsbacher, M.A., Goldsmith, H.H. & Robertson, R.R.W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion*, 6, pp. 89-111.

Gerrig, R.J. (1988). Text comprehension. In R.J. Sternberg & E.E. Smith (Eds.), *The psychology of human thought*, pp. 242-266. Cambridge: Cambridge University Press.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gilhooly, K.J. (1988). *Thinking: directed, undirected and creative*. London: Academic Press.

Gilhooly, K.J., Logie, R.H., Wetherick, N.E. & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory and Cognition*, 21, pp. 115-124.



- Glenberg, A.M. & Langston, W.E. (1992). Comprehension of illustrated text: pictures help to build mental models. *Journal of Memory and Language*, 31, pp. 129-151.
- Glenberg, A.M., Meyer, M. & Lindberg, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26, pp. 69-83.
- Graesser, A.C., Gordon, S.E. & Sawyer, J.D. (1979). Recognition memory for typical and atypical actions in scripted activities: tests of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behaviour*, 18, pp. 319-332.
- Hagert, G. (1984). Modeling mental models: experiments in cognitive modeling of spatial reasoning. In T. O'Shea (Ed.), *Advances in artificial intelligence*, (pp. 389-398). Amsterdam: North-Holland.
- Hammerton, M. & Altham, P.M.E. (1971). A non-parametric alternative to  $d'$ . *Nature*, 234, pp. 487-488.
- Hampson, P.J. & Morris, P.E. (1979). Cyclical processing: a framework for imagery research. *Journal of Mental Imagery*, 3(1-2), pp. 11-22.
- Hanley, J.R., Young, A.W. & Pearson, N.A. (1991). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology*, 43A, pp. 101-125.
- Hayes, P.J. (1979). The naïve physics manifesto. In D. Michie (Ed.), *Expert systems in the micro-electronic age*, pp. 242-270. Edinburgh: Edinburgh University Press.
- Hayes, P.J. (1985). The second naive physics manifesto. In J.R. Hobbs & R.C. Moore (Eds.), *Formal theories of the commonsense world*, pp. 1-20. Norwood, NJ: Ablex.
- Henle, M. (1962). The relation between logic and thinking. *Psychological Review*, 69, pp. 366-378.
- Herman, J.F., Cachuela, G.M. & Heins, J.A. (1987). Children's and adults' long-term memory for spatial locations over an extended time period. *Developmental Psychology*, 23, pp. 509-513.
- Herman, J.F., Kail, R.V. & Siegel, A.W. (1979). Cognitive maps of a college campus: a new look at freshman orientation. *Bulletin of the Psychonomic Society*, 13, pp. 183-186.
- Hilgard, E.R., Leary, D.E. & McGuire, G.R. (1991). The history of psychology: a survey and critical assessment. *Annual Review of Psychology*, 42, pp. 79-107.
- Hintzman, D.L. (1978). *The psychology of learning and memory*. W.H. Freeman: San Francisco.
- Hintzman, D.L. (1990). Human learning and memory: connections and dissociations. *Annual Review of Psychology*, 41, pp. 109-139.
- Holland, J.H., Holyoak, K.J., Nisbett, R.E. & Thagard, P.R. (1986). *Induction: processes of inference, learning and discovery*. Cambridge, MA: MIT Press.
- Holyoak, K.J. (1993). Thinking. *Annual Review of Psychology*, 44, pp. 265-315.
- Horton, D.L. & Mills, C.B. (1984). Human learning and memory. *Annual Review of Psychology*, 35, pp. 361-394.



Howell, D.C. (1992). *Statistical methods for psychology* (3rd. ed.). Belmont, CA: Duxberry Press.

Hunt, E. (1987). A cognitive model of individual differences, with an application to attention. In S.H. Irvine & S.E. Newstead (Eds.), *Intelligence and cognition: contemporary frames of reference*, pp. 177-215. Dordrecht: Martinus Nijhoff.

Hunt, E. (1989). Cognitive science: definition, status and questions. *Annual Review of Psychology*, 40, pp. 603-29.

Hunt, R.R. & Einstein, G.O. (1980). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behaviour*, 20, pp. 497-514.

Hunter, I.M.L. (1957). The solving of three term series problems. *British Journal of Psychology*, 48, pp. 286-296.

Intraub, H. (1979). The role of implicit naming in pictorial encoding. *Journal of Experimental Psychology: Human Perception and Performance*, 5, pp. 78-87.

Jacoby, L.L. & Craik, F.I.M. (1979). Effects of elaboration of processing at encoding and retrieval: trace distinctiveness and recovery of initial context. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 1-21. Hillsdale, NJ: Lawrence Erlbaum Associates.

James, W. (1890). *Principles of psychology*. New York: Holt.

Johnson-Laird, P.N. (1980). Mental models in cognitive science. *Cognitive Science*, 4, pp. 71-115.

Johnson-Laird, P.N. (1981). Comprehension as the construction of mental models. *Philosophical Transactions of the Royal Society of London, Series B*, 295(1077), pp. 353-374.

Johnson-Laird, P.N. (1983). *Mental models: towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (1986). Reasoning without logic. In T. Myers, K. Brown & B. McGonigle (Eds.), *Reasoning and discourse processes*, pp. 13-49. London: Academic Press.

Johnson-Laird, P.N. (1988a). A computational analysis of consciousness. In A.J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science*, pp. 357-368. Oxford: Oxford Science Publications.

Johnson-Laird, P.N. (1988b). *The computer and the mind: an introduction to cognitive science*. London: Fontana Press.

Johnson-Laird, P.N. (1989). Mental models. In M.I. Posner (Ed.), *Foundations of Cognitive Science*, pp. 469-499. Cambridge, MA: MIT Press.

Johnson-Laird, P.N. (1993). *Human and machine thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P.N. (1994). Mental models and probabilistic thinking. *Cognition*, 50, pp. 189-209.

Johnson-Laird, P.N. & Byrne, R.M.J. (1989). Only reasoning. *Journal of Memory and Language*, 28, pp. 313-330.



- Johnson-Laird, P.N. & Byrne, R.M.J. (1990). Meta-logical problems: knights, knaves and Rips. *Cognition*, 36, pp. 69-84.
- Johnson-Laird, P.N. & Byrne, R.M.J. (1991). *Deduction*. Hove: Lawrence Erlbaum Associates.
- Johnson-Laird, P.N. & Byrne, R.M.J. (1993). Précis of *Deduction*. *Behavioral and Brain Sciences*, 16, pp. 323-380.
- Johnson-Laird, P.N., Byrne, R.M.J. & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99, pp. 418-439.
- Johnson-Laird, P.N., Byrne, R.M.J. & Tabossi, P. (1989). Reasoning by model: the case of multiple quantification. *Psychological Review*, 96, pp. 658-673.
- Johnson-Laird, P.N., Gibbs, G. & de Mowbray, J. (1978). Meaning, amount of processing and memory for words. *Memory and Cognition*, 6, pp. 372-375.
- Johnson-Laird, P.N. & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology*, 10, pp. 64-99.
- Jones, G.V. (1978). Recognition failure and dual mechanisms in recall. *Psychological Review*, 85, pp. 464-469.
- Jones, G.V. (1982). Tests of the dual-mechanism theory of recall. *Acta Psychologica*, 50, pp. 61-72.
- Jones, G.V. (1987). Independence and exclusivity among psychological processes: implications for the structure of recall. *Psychological Review*, 94, pp. 229-235.
- Jordan, M.I. & Rosenbaum, D.A. (1989). Action. In M.I. Posner (Ed.), *Foundations of cognitive science*, pp. 727-767. Cambridge, MA: MIT Press/Bradford Books.
- Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99, pp. 122-149.
- Kelso, J.A.S. (Ed.). (1982). *Human motor behaviour: an introduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, 3, pp. 87-108.
- Kintsch, W. (1988). The role of knowledge in discourse-comprehension: a construction-integration model. *Psychological Review*, 95, pp. 163-182.
- Kintsch, W. (1994). Text comprehension, memory and learning. *American Psychologist*, 49, pp. 294-303.
- Kintsch, W., Mandel, T.S. & Kozminsky, E. (1977). Summarizing scrambled stories. *Memory and Cognition*, 5, pp. 547-552.
- Kirby, K.N. & Kosslyn, S.M. (1990). Thinking visually. *Mind and Language*, 5, pp. 324-341.
- Kosslyn, S.M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.



Kosslyn, S.M. (1981). The medium and the message in mental imagery: a theory. *Psychological Review*, 88, pp. 44-66.

Kosslyn, S.M. (1984). Mental representation. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honour of Gordon Bower*, pp. 91-117. San Francisco, CA: W.H. Freeman.

Kosslyn, S.M. (1991). A cognitive neuroscience of visual cognition: further developments. In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition*, pp. 351-381. Amsterdam: North-Holland.

Kosslyn, S.M., Ball, T.M. & Reiser, B.J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4, pp. 47-60.

Kosslyn, S.M., Cave, C.B., Provost, D.A. & von Gierke, S.M. (1988). Sequential processes in image generation. *Cognitive Psychology*, 20, pp. 319-343.

Kosslyn, S.M. & Pomerantz, J.R. (1977). Imagery, propositions and the form of internal representations. *Cognitive Psychology*, 9, pp. 52-76.

Kosslyn, S.M., Seger, C., Pani, J.R. & Hillger, L.A. (1990). When is imagery used in everyday life? A diary study. *Journal of Mental Imagery*, 14, pp. 131-152.

Kucera, H. & Francis, W.N. (1967). *Computational analysis of present-day american english*. Providence: Brown University Press.

Kulhavy, R.W., Stock, W.A., Peterson, S.E., Pridemore, D.R. & Klein, J.D. (1992). Using maps to retrieve text: a test of conjoint retention. *Contemporary Educational Psychology*, 17, pp. 56-70.

Kulhavy, R.W., Stock, W.A., Verdi, M.P., Rittschof, K.A. & Savenye, W. (1993). Why maps improve memory for text: the influence of structural information on working memory operations. *European Journal of Cognitive Psychology*, 5, pp. 375-392.

Kulhavy, R.W., Stock, W.A., Woodard, K.A. & Haygood, R.C. (1993). Comparing elaboration and dual coding theories: the case of maps and text. *American Journal of Psychology*, 106, pp. 483-498.

Kulhavy, R.W., Woodard, K.A., Haygood, R.C. & Webb, J.M. (1993). Using maps to remember text: an instructional analysis. *British Journal of Educational Psychology*, 63, pp. 161-169.

Kyllonen, P.C. & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, pp. 389-433.

Langer, P., Keenan, V. & Bergman, J. (1993). Contributions of different feedback assistance to text memorial representation. *Bulletin of the Psychonomic Society*, 31, pp. 209-212.

Langer, P., Keenan, V. & Cumbo, K. (1992). The effects of text version and feedback type on memorial representations. *Bulletin of the Psychonomic Society*, 30, pp. 373-376.

Langer, P., Keenan, V. & Nelson, S. (1991). Influence of feedback type on comprehension with two variants of a text. *Bulletin of the Psychonomic Society*, 29, pp. 348-350.



- Langer, P., Keenan, V. & Nelson, S. (1991). Sentence-order feedback during processing of sequential or spatial texts. *Bulletin of the Psychonomic Society*, 29, pp. 31-32.
- Le Voi, M. (1984). *A model of recognition failure*. (HCRL Technical Report 9). The Open University.
- Leahey, T.H. (1991). *A history of modern psychology*. London: Prentice-Hall.
- Leahey, T.H. (1992). The mythical revolutions of American psychology. *American Psychologist*, 47, pp. 308-318.
- Lee, D.N. (1980). The optic flow field: the foundation of vision. *Philosophical Transactions of the Royal Society of London, Series B*, 290, pp. 169-179.
- Levelt, W.J.M. (1984). Some perceptual limitations on talking about space. In A.J. van Doorn, W.A. de Grind & J.J. Koenderink (Eds.), *Limits in perception*, pp. 323-358. Utrecht: VNU Science Press.
- Levine, M., Jankovic, I. & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, 111, pp. 157-175.
- Logie, R.H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology*, 38A, pp. 229-247.
- Logie, R.H. (1989). Characteristics of visual short-term memory. *European Journal of Cognitive Psychology*, 1, pp. 275-284.
- Logie, R.H. (1990). Interference with visual short-term memory. *Acta Psychologica*, 75, pp. 55-74.
- Logie, R.H. & Marchetti, C. (1991). Visuo-spatial working memory: visual, spatial or central executive? In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition*, pp. 105-115. Amsterdam: North-Holland.
- Lycan, W.G. (Ed.). (1990). *Mind and cognition: a reader*. Oxford: Blackwell.
- Mandler, G. (1980). Recognizing: the judgement of previous occurrence. *Psychological Review*, 87, pp. 252-271.
- Mani, K. & Johnson-Laird, P.N. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, 10, pp. 181-187.
- Manktelow, K. & Over, D. (1992). Obligation, permission and mental models. In Y. Rogers, A. Rutherford & P.A. Bibby (Eds.), *Models in the mind: theory, perspective and application*, pp. 249-266. London: Academic Press.
- Marascuilo, L.A. & McSweeney, M. (1967). Nonparametric post hoc comparisons for trend. *Psychological Bulletin*, 67, pp. 401-412.
- Marks, D.F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64, pp. 17-24.



- Marks, D.F. (1989). Images as mental models. In A.F. Bennett & K.M. McKonkey (Eds.), *Cognition in individual and social contexts*, pp. 37-46. Amsterdam: Elsevier.
- Marks, D.F. (1990). On the relationship between imagery, body and mind. In P.J. Hampson, D.F. Marks & J.T.E. Richardson (Eds.), *Imagery: current developments*, pp. 1-38. London: Routledge.
- Marr, D. (1980). Visual information processing. *Philosophical Transactions of the Royal Society of London, Series B*, 290, pp. 209-218.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.
- Marschark, M. (1985). Imagery and organization in the recall of prose. *Journal of Memory and Language*, 24, pp. 734-745.
- Marschark, M. & Surian, L. (1989). Why does imagery improve memory? *European Journal of Cognitive Psychology*, , pp. .
- McClelland, J.L. & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: part 1. An account of basic findings. *Psychological Review*, 88, pp. 375-407.
- McGuinness, C. (1989). Visual imagery: the question of representation. *The Irish Journal of Psychology*, 10, pp. 188-200.
- McNamara, T.P. (1991). Memory's view of space. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 27. pp. 147-186. San Diego: Academic Press.
- McNamara, T.P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, 99, pp. 650-662.
- McNamara, T.P., Halpin, J.A. & Hardy, J.K. (1992a). The representation and integration in memory of spatial and nonspatial information. *Memory and Cognition*, 20, pp. 519-532.
- McNamara, T.P., Halpin, J.A. & Hardy, J.K. (1992b). Spatial and temporal contributions to the structure of spatial memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, pp. 555-564.
- Meddis, R. (1984). *Statistics using ranks: a unified approach*. Oxford: Blackwell.
- Miller, G.A. & Johnson-Laird, P.N. (1976). *Language and Perception*. Cambridge: Cambridge University Press.
- Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision*, pp. 211-277. New York: McGraw-Hill.
- Montgomery, H. (1988). Mental models and problem solving: three challenges to a theory of restructuring and insight. *Scandinavian Journal of Psychology*, 29, pp. 85-94.
- Morra, S. (1989). Developmental differences in the use of verbatim versus spatial representations in the recall of spatial descriptions: a probabilistic model and an experimental analysis. *Journal of Memory and Language*, 28, pp. 37-55.



- Morra, S., Moizo, C. & Scopesi, A. (1988). Working memory (or the M operator) and the planning of children's drawings. *Journal of Experimental Child Psychology*, 46, pp. 41-73.
- Morra, S., Pascual-Leone, J., Johnson, J. & Baillargeon, R. (1991). Understanding spatial descriptions: test of a mental capacity model. In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition*, pp. 241-254. Amsterdam: North-Holland.
- Morris, C.D., Bransford, J.D. & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behaviour*, 16, pp. 519-533.
- Morris, P.E. & Reid, R.L. (1972). Canadian and British ratings of the imagery values of words. *British Journal of Psychology*, 63, pp. 163-154.
- Morris, P.E. & Reid, R.L. (1975). The role of stimulus and pair imagery in paired-associate learning. *British Journal of Psychology*, 66, pp. 153-156.
- Morris, P.E. & Stevens, R. (1974). Linking images and free recall. *Journal of Verbal Learning and Verbal Behavior*, 13, pp. 310-315.
- Morrow, D.G., Bower, G.H. & Greenspan, S.L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, 28, pp. 292-312.
- Morrow, D.G., Bower, G.H. & Greenspan, S.L. (1990). Situation-based inferences during narrative comprehension. In A.C. Graesser & G.H. Bower (Eds.), *Inferences and text comprehension, The psychology of learning and motivation: advances in research and theory, Volume 25*. pp. 123-135. San Diego: Academic Press.
- Morrow, D.G., Greenspan, S.L. & Bower, G.H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language*, 26, pp. 165-187.
- Nelson, D.L. (1979). Remembering pictures and words: appearance, significance, and name. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 45-76. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Newell, A. (1990). *Unified theories of cognition*. Harvard, MA: Harvard University Press.
- Newell, A. (1992). Précis of *Unified theories of cognition*. *Behavioral and Brain Sciences*, 15, pp. 425-492.
- Newstead, S.E. & Evans, J.S.B.T. (1993). Mental models as an explanation of belief bias effects in syllogistic reasoning. *Cognition*, 46, pp. 93-97.
- Noordman, L.G.M. & Vonk, W. (1992). Reader's knowledge and the control of inferences in reading. *Language and Cognitive Processes*, 7, pp. 373-391.
- Norman, D.A. (1983). Some observations on mental models. In D. Gentner & A.L. Stevens (Eds.), *Mental models*, pp. 7-14. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Norman, D.A. & Rumelhart, D.E. (1975). Memory and knowledge. In D.A. Norman & D.E. Rumelhart (Eds.), *Explorations in cognition*, pp. 1-32. San Francisco, CA: W.H. Freeman.
- O'Brien, E.J. & Albrecht, J.E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, pp. 777-784.



- Oakhill, J.V. & Garnham, A. (1993). On theories of belief bias in syllogistic reasoning. *Cognition*, 46, pp. 87-92.
- Oakhill, J.V., Garnham, A. & Johnson-Laird, P.N. (1990). Belief bias effects in syllogistic reasoning. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (Eds.), *Lines of thinking: reflections on the psychology of thought, Volume 1: representation, reasoning, analogy and decision making*. pp. Chichester: John Wiley.
- Oakhill, J.V. & Johnson-Laird, P.N. (1984). Representation of spatial description in working memory. *Current Psychological Research and Reviews*, 3, pp. 52-62.
- Oakhill, J.V. & Johnson-Laird, P.N. (1985a). The effects of belief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology*, 37A, pp. 553-569.
- Oakhill, J.V. & Johnson-Laird, P.N. (1985b). Rationality, memory and the search for counterexamples. *Cognition*, 20, pp. 79-94.
- Oakhill, J.V., Johnson-Laird, P.N. & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, 31, pp. 117-140.
- Ormerod, T.C., Manktelow, K.I. & Jones, G.V. (1993). Reasoning with types of conditional: biases and mental models. *Quarterly Journal of Experimental Psychology*, 46A, pp. 653-677.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Paivio, A. (1986). *Mental representation: a dual coding approach*. New York: Oxford University Press.
- Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology*, 45, pp. 255-287.
- Paivio, A., Yuille, J.C. & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs*, 76(1, Pt. 2), pp. 1-25.
- Palmer, S.E. (1975). Visual perception and world knowledge: notes on a model of sensory-cognitive interaction. In D.A. Norman & D.E. Rumelhart (Eds.), *Explorations in cognition*, pp. 279-307. San Francisco: W.H. Freeman.
- Palmer, S.E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization*, pp. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Palmer, S.E. & Kimchi, R. (1986). The information processing approach to cognition. In T.J. Knapp & R.C. Robertson (Eds.), *Approaches to cognition: constraints and controversies*, pp. 37-77. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Payne, S.J. (1988). Methods and mental models in theories of cognitive skill. In J. Self (Ed.), *Artificial intelligence and human learning: intelligent computer aided instruction*, pp. London: Chapman and Hall.
- Payne, S.J. (1991). A descriptive study of mental models. *Behaviour and Information Technology*, 10, pp. 3-21.



- Payne, S.J. (1993). Memory for mental models of spatial descriptions: an episodic construction trace account. *Memory and Cognition*, 21, pp. 591-603.
- Perrig, W. & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language*, 24, pp. 503-518.
- Piaget, J. & Inhelder, B. (1958). *The growth of logical thinking from childhood to adolescence*. New York, NY: Basic Books.
- Pichert, J.W. & Anderson, R.C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, 69, pp. 309-315.
- Pinker, S. & Prince, A. (1988). On language and connectionism analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, pp. 73-193.
- Potts, G.R. (1972). Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behaviour*, 11, pp. 727-740.
- Pylyshyn, Z.W. (1973). What the mind's eye tells the mind's brain. *Psychological Bulletin*, 80, pp. 1-24.
- Pylyshyn, Z.W. (1984). *Computation and cognition: towards a foundation for cognitive science*. Cambridge, MA: MIT Press.
- Pylyshyn, Z.W. (1989). Computing in cognitive science. In M.I. Posner (Ed.), *Foundations of cognitive science*, pp. 51-91. Cambridge, MA: MIT Press/Bradford Books.
- Quinn, G. (1991). Encoding and maintenance of information in visual working memory. In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition*, pp. 95-104. Amsterdam: North-Holland.
- Radvansky, G.A., Spieler, D.H. & Zacks, R.T. (1993). Mental model organization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, pp. 95-114.
- Radvansky, G.A. & Zacks, R.T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, pp. 940-953.
- Reder, L.M. (1979). The role of elaborations in memory for prose. *Cognitive Psychology*, 11, pp. 221-234.
- Reder, L.M. (1980). The role of elaboration in the comprehension and retention of prose: a critical review. *Review of Educational Research*, 50, pp. 5-53.
- Reed, S.K. (1974). Structural descriptions and the limitations of visual images. *Memory and Cognition*, 2, pp. 329-336.
- Reed, S.K., Ernst, G.W. & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, 6, pp. 436-450.
- Richardson, J.T.E. (1987). The role of mental imagery in models of transitive inference. *British Journal of Psychology*, 78, pp. 189-203.



- Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90, pp. 38-71.
- Rips, L.J. (1986). Mental muddles. In M. Brand & R.M. Harnish (Eds.), *The representation of knowledge and belief*, pp. Tuscon: University of Arizona Press.
- Rips, L.J. (1990). Paralogical reasoning. *Cognition*, 36, pp. 291-314.
- Roediger, H.L. (1990). Implicit memory: retention without remembering. *American Psychologist*, 45, pp. 1043-1056.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, pp. 192-233.
- Rosch, E. & Mervis, C.B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, 7, pp. 573-605.
- Rosenbaum, D.A. (1984). Planning and control of motor movements. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honour of Gordon Bower*, pp. 219-233. San Francisco: W.H. Freeman.
- Roth, I. & Frisby, J.P. (1986). *Perception and representation: a cognitive approach*. Milton Keynes: Open University Press.
- Rumelhart, D.E. (1989). The architecture of mind: a connectionist approach. In M.I. Posner (Ed.), *Foundations of cognitive science*, pp. 133-159. Cambridge, MA: MIT Press/Bradford Books.
- Rumelhart, D.E. & McClelland, J.L. (1985). Levels indeed! A response to Broadbent. *Journal of Experimental Psychology: General*, 114, pp. 193-197.
- Rumelhart, D.E. & Norman, D.A. (1983). *Representation in memory*. (CHIP Technical Report 116). Centre for Human Information Processing, University of San Diego.
- Rumelhart, D.E., Smolensky, P., McClelland, J.L. & Hinton, G.E. (1986). Schemata and sequential thought processes in PDP models. In J.L. McClelland & D.E. Rumelhart (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition, Volume 2: Psychological and biological models*. pp. 7-57. Cambridge, MA: MIT Press/Bradford Books.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Schaeken, W., Johnson-Laird, P.N. & d'Ydewalle, G. (1993). *Mental models and temporal reasoning*. (Psychological Reports No. 149). University of Leuven.
- Schank, R.C. (1975). *Conceptual information processing*. New York: North-Holland.
- Schank, R.C. (1982). *Dynamic memory: a theory of reminding and learning in computers and people*. Cambridge: Cambridge University Press.
- Schank, R.C. (1982). Reminding and memory organization. In W.G. Lehnert & M.H. Ringle (Eds.), *Strategies for natural language processing*, pp. Hillsdale, NJ: Lawrence Erlbaum Associates.



- Schank, R.C. & Abelson, R.P. (1977). *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schmalhofer, F. & Glavanov, D. (1986). Three components of understanding a programmer's manual: verbatim, propositional and situational representations. *Journal of Memory and Language*, 25, pp. 279-294.
- Selfridge, O.G. (1959). Pandemonium: a paradigm for learning. In D.V. Blake & A.M. Uttley (Eds.), *Symposium on the mechanization of thought processes*, (pp. 511-529). London: HMSO.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Shepard, R.N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behaviour*, 6, pp. 156-163.
- Shepard, R.N. (1975). Form, formation and transformation of internal representations. In R.L. Solso (Ed.), *Information processing and cognition: the Loyola symposium*, pp. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shepard, R.N. & Chipman, S. (1970). Second order isomorphism of internal representations: shapes of states. *Cognitive Psychology*, 1, pp. 1-17.
- Shepard, R.N. & Cooper, L.A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- Shepard, R.N. & Feng, C. (1972). A chronometric study of paper folding. *Cognitive Psychology*, 3, pp. 228-243.
- Shepard, R.N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, pp. 701-703.
- Shepard, R.N. & Podgorny, P. (1978). Cognitive processes that resemble perceptual processes. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes*, pp. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Siegel, S. & Castellan, N.J., Jr. (1988). *Nonparametric statistics for the behavioural sciences* (2nd ed.). New York: McGraw-Hill.
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioural and Brain Sciences*, 11, pp. 1-74.
- Smyth, M.M., Morris, P.E., Levy, P. & Ellis, A.W. (1988). *Cognition in action*. London: Lawrence Erlbaum Associates.
- Speelman, C.P. & Kirsner, K. (1990). The representation of text-based and situation-based information in discourse comprehension. *Journal of Memory and Language*, 29, pp. 119-132.
- Tardieu, H., Ehrlich, M.-F. & Gyselinck, V. (1992). Levels of representation and domain-specific knowledge in comprehension of scientific texts. *Language and Cognitive Processes*, 7, pp. 335-351.
- Taylor, H.A. & Tversky, B. (1992a). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31, pp. 261-292.



Taylor, H.A. & Tversky, B. (1992b). Descriptions and depictions of environments. *Memory and Cognition*, 20, pp. 483-496.

Thorndyke, P.W. (1984). Applications of schema theory in cognitive research. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honour of Gordon Bower*, pp. 167-191. San Francisco: W.H. Freeman.

Thorndyke, P.W. & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14, pp. 560-589.

Tolman, E.C. (1932). *Purposive behaviour in animals and men*. New York: Appleton-Century-Crofts.

Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, pp. 189-209.

Toms, M., Morris, N. & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology*, 46A, pp. 679-699.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Ed.), *Organization of memory*, pp. New York: Academic Press.

Tulving, E. (1979). Relation between encoding specificity and levels of processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*, pp. 405-428. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tulving, E. & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, pp. 352-373.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, pp. 327-352.

Tversky, B. (1991). Spatial mental models. In G.H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, 27, pp. 109-145. San Diego: Academic Press.

van Dijk, T.A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vandierendonck, A. & de Vooght, G. (1992). *Is reasoning with time concepts based on a spatialized representation of time?* (Reports of the Department of General Psychology No. 2). University of Ghent.

Vandierendonck, A. & de Vooght, G. (1993). *Shadowing and generation of random time intervals: a pure central executive interference task*. (Reports of the Department of General Psychology No. 8). University of Ghent.

Waddell, K.J. & Rogoff, B. (1987). Contextual organization and intentionality in adult's spatial memory. *Developmental Psychology*, 23, pp. 509-513.

Watkins, M.J. & Gardiner, J.M. (1979). An appreciation of generate-recognize theory of recall. *Journal of Verbal Learning and Verbal Behaviour*, 18, pp. 687-704.

Watson, J.B. (1913). Psychology as the behaviourist views it. *Psychological Review*, 20, pp. 158-177.



Weldon, M.S. & Roediger, H.L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory and Cognition*, 15, pp. 269-280.

Wilson, S.G., Rinck, M., McNamara, T.P., Bower, G.H. & Morrow, D.G. (1993). Mental models and narrative comprehension: some qualifications. *Journal of Memory and Language*, 32, pp. 141-154.

Yuille, J.C. & Catchpole, M.J. (1977). The role of imagery in models of cognition. *Journal of Mental Imagery*, 1, pp. 171-180.